I. R. IRAN

ISSN: 1728-1431
e-ISSN: 1735-9244

International Journal of Engineering

Journal Homepage: www.ije.ir

MERC

TRANSACTIONS A: Basics

Volume 34, Number 04, April 2021

Materials and Energy Research Center

In the Name of God

INTERNATIONAL JOURNAL OF ENGINEERING
Transactions A: Basics

DISCLAIMER

The publication of papers in International Journal of Engineering does not imply that the editorial board, reviewers or publisher accept, approve or endorse the data and conclusions of authors.

International **J**ournal of **E**ngineering *Transactions A: Basics* (ISSN 1728-1431) (EISSN 1735-9244)
International **J**ournal of **E**ngineering *Transactions B: Applications* (ISSN 1728-144X) (EISSN 1735-9244)
International **J**ournal of **E**ngineering *Transactions C: Aspects* (ISSN 2423-7167) (EISSN 1735-9244)
Web Sites: www.ije.ir & www.ijeir.info E-mails: ije.editor8@gmail.com, Tel: (+9821) 88771578, Fax: (+9821) 88773352
Materials and Energy Research Center (MERC)

ISSN: 1728-1431, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 4, April 2021

CONTENTS:

ISSN: 1728-1431, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 4, April 2021

Industrial Engineering

ISSN: 1728-1431, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 4, April 2021

Material Engineering

Mechanical Engineering

ISSN: 1728-1431, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 4, April 2021

Mining Engineering

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 756-762

International Journal of Engineering

Journal Homepage: www.ije.ir

Photo-degradation of P-Nitro Toluene using Modified Bentonite Based Nano-TiO_2 Photocatalyst in Aqueous Solution

M. Rostami, A. Hassani Joshaghani*, H. Mazaheri, A. Shokri

Department of Chemical Engineering, Arak Branch, Islamic Azad University, Arak, Iran

P A P E R I N F O

Paper history:
Received 21 November 2020
Received in revised form 23 December 2020
Accepted 14 January 2021

Keywords:
Titanium Dioxide
Bentonite
Advanced Oxidation
P-Nitro Toluene

A B S T R A C T

In recent decades, Iran has been facing severe water deficiency. In all countries, industrial plants are the most water-consuming sectors; thus, industrial wastewater treatment is always a essential subject. Nitro-Toluene derivatives are extensively used in industries, especially the military industry, which itself has an abundant share in industrial wastewater contamination. These compounds are extremely dangerous for living beings and can have irreparable effects, so eradication of them in industrial wastewater is necessary. Photocatalytic processes are one of the particular approaches in industrial wastewater treatment from the advanced oxidation processes subdivision. One of the prominent and most widely used photocatalysts in this process is Titanium Dioxide (TiO_2) . This research aims at the investigations for the modification of TiO_2/Bentonite (TB) catalysts for attaining more economical saving and degradation stabilization conditions. To achieve this goal, the Bentonite and TiO_2 photocatalyst was synthesized by a co-precipitation procedure, and its catalytic activity on Para Nitro-Toluene (PNT) degradation was examined. The designed TB photocatalyst is made of 5, 10 and 20 % of TB. A suspension reactor and the spectrophotometry was applied for specifying the extent of the degradation. Characterization of modified catalyst was conducted by scanning electron microscopy (SEM), X-ray diffraction (XRD), Fourier transform infrared (FT-IR) spectroscopy and energy dispersive X-ray (EDX). The results highlight that with increasing TiO_2 percent, degradation rate augmented, and the highest degradation was attained for TB 20% at 59%. However, Under the same conditions, for pure TiO_2, the degradation rate is 64%, but with more TiO_2 consumption and time. Finally, in order to further confirm the extent of the degradation, chemical oxygen demand (COD) test was performed on the TA 20 sample. The results showed that about 53% of PNT has been converted to minerals.

doi:*10.5829/ije.2021.34.04a.01*

1. INTRODUCTION

Due to the global shortage of water resources and their continuous depletion, along with the growth of human communities and the manufacturing of a wide range of chemicals in various industries, the need to develop efficient methods in water and wastewater treatment have become more considerable and the wastewater reuse will be vital [1, 2].

Aromatic compounds are one of the most polluting objects in the world and also in Iran. Some organic compounds are disposed by biological methods in effluents, but despite their dominance, they also have inferiorities such as incompetence in the presence of biodegradable organic compounds and the production of noxious intermediates. Types of compounds It is essential to use novel processes such as advanced oxidation that can convert non-biodegradable molecules into molecules that can be degraded by microorganisms or into inert molecules such as water, carbon dioxide, nitrogen, and so on [3, 4].

Toluene and its nitrate derivatives are widely used in differing industries. These compounds are employed in defense industries, chemicals, rubber, insecticides, textile and paper industries, paint industries, and so on. These compounds are often recognized as environmental pollutants, especially in industrial

*Corresponding Author Email: a-hasani@iau-arak.ac.ir (A. Hassani Joshaghani)

Please cite this article as: M. Rostami, A. Hassani Joshaghani, H. Mazaheri, A. Shokri, Photo-degradation of P-Nitro Toluene Using Modified Bentonite Based Nano-TiO_2 Photocatalyst in Aqueous Solution, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 756-762

wastewater treatment. Nitro-Toluene derivatives include 2-nitro-toluene, 3-nitro-toluene, and 4-nitro-toluene. These aromatic derivatives are toxic, and their demolition is mandatory. It is noteworthy that 4-nitro-toluene or the same PNT is the most abundant derivative available [5]. Many procedures including preparation of pesticides, textiles and paper use nitroaromatics, which are damaging to human and environment if discharged into water lacking proper treatment. The p-nitrotoluene (PNT) is one of the public nitroaromatics and studies exhibited that it is a suspected hormone disrupter. So, some studies have been paid to the degradation of PNT in wastewater.

One of the effective methods to degradation of PNT contaminants is the use of advanced oxidation processes, which were discussed below.

Advanced oxidation processes are generally referred to processes in which a strong oxidizing agent such as oxygenated water, ozone, and a catalyst such as titanium dioxide, zinc oxide, iron, and manganese are applied in the presence or absence of UV radiation [6,7]. These processes rely on the production of hydroxyl radicals with high oxidation activity, which convert organic chemical pollutants into minerals that are most efficient in oxidizing of durable organic compounds. When the energy of a photon is equal to or greater than of semiconductor, it leads to excitation of the electron energy gap (Eg) from the capacitance band to the conduction band, which due to the excitation of the electron generate a gap in capacitance band. Excited electrons can directly or indirectly create hydroxyl radicals, in which these radicals alter the organic matter to minerals. The use of titanium dioxide photocatalyst is recognized as an efficient method of advanced oxidation methods [8,9].

In this study, TiO_2 was used as a photocatalyst. TiO_2 exists in the three main phases of Rutile, Anatase, and Brucite. Of these three phases, the Rutile phase is the most reliable, and the other two phases, namely Anatase and Brucite, are transformed into the Rutile phase by gaining heat. Spatially, the Rutile and Anatase phases are tetragonal, and the Brucite has an orthorhombic shape. However, for some applications such as conventional solution filtration, the crystal phase is not required. The crystalline phase is necessary when a particular application, such as photocatalytic or semiconductor, is considered. [10,11].

Different materials can be combined with TiO_2 as a base to serve different purposes. Therefore, in this study, in order to increase the contact level and reduce the cost, bentonite was used as a cheap and available material. Bentonite is a clay material composed of swollen minerals, which are mainly Montmorillonite and, to a small extent, is Bidellite. Most bentonites are created by weathering and alteration of volcanic ash, often in the presence of water, and their source rocks are mostly basalt. Volcanic ash decomposition is mainly carried out in saline and swampy environments, and as we move away from the volcano, the thickness of bentonite decreases. Bentonite has an alumina silicate structure and belongs to the three-layer silicates classes. Bentonite has two quadrilateral layers and one octahedral layer [12].

In recent years, studies have been done on the combination of TiO_2 with other materials, which have been referred to in the following : Nasirian et al. [13] reported the photocatalytic behaviors of the Fe_2O_3 / TiO_2 on the degradation of ordinary dyes. In their research, the degradation of Congo Red (CR) and Methyl Orange (MO) was investigated, and the results revealed that by using 0.01 wt% Fe_2O_3/ TiO_2 as a photocatalyst, 62% of MO and 46.8% of CR could be degraded. Also, the optimum temperature in this process was 300°C, where the highest color degradation occurs [13].

Sethau et al. [14] studied the elimination of the methylene blue dye using a TiO_2 zeolite photocatalyst. In their research, after describing the synthesizing procedure, the performance of the synthesized zeolite TiO_2 photocatalyst is exposed in comparison with the pure TiO_2 which highlights the higher performance and greater efficiency of the synthesized photocatalyst [14].

Hosseini and et al. [15] presented a study on photocatalytic degradation of 4, 2-dichlorophenol using TiO_2 nanoparticles and also Co/TiO_2 containing mixed matrix membranes. They synthesized TiO_2 with different amounts of cobalt by the sol-gel method. The best combination occurred with 1.34 molar percent of Co, which had the highest degradation rate. They also studied the TiO_2 degradation with mixed membranes and also applied PES membranes and found that adding 1 wt% of Co/TiO_2 (1.34%) to Polyethersulfone (PES) had a greater passing flux. They also reported the best method for the separation of 2, 4 dichlorophenols [15]. Also, Gharibshahian [16] investigated the influence of the concentration of Polyvinyl Alcohol on the progress kinetics of $KTiOPO_4$ nanoparticles produced by the co-precipitation Method. Preparation, identification and optical features of $LaFeO_3$ nanoparticles through sol-gel combustion method was explored by Theingi et al. [17].

Aby et al. [18] published an article on how to make a nano-silver and zinc (Ag/ZnO) photocatalyst and demonstrated its elevated efficiency. The effect of photocatalyst degradation on organic pollutants such as some dyes was investigated, and it was concluded that the efficacy of ZnO <Ag / ZnO (% 1wt) <Ag / ZnO (4 wt%) [18].

As has been observed, many studies have been done, but studies are at the boundary of knowledge and have the potential to be expanded. Therefore, the present study investigates the combination of TiO_2 and bentonite to reduce photocatalyst cost and Finally, the efficiency of the photocatalyst made on PNT degradation was evaluated.

2. EXPERIMENTAL

2. 1. Materials and Methods

The used materials of this study involve PNT manufactured by Merck Company, TiO_2 produced by the US Research Nano Materials Company (APS of 20 nm and BET surface area of 200 m^2/g). Bentonite soil was purchased from the Chinese DAE JUNG Company. Agilent 8453 spectrophotometers, KBr-PerkinElmer Fourier transform infrared (FT-IR) analyzer, PW1730 Philips-Diffractometer X-ray diffraction (XRD) analyzer, Phenom ProX Scanning electron microscopy (SEM) and Spectroscopy and energy dispersive X-ray (EDX) for characterizing the synthesized photocatalyst and all the experimental activities were performed in a home-made laboratory pilot (semi CREC reactor with capacity of 1.4 liters and height of 41 cm) [19,20].

2. 2. Synthesis of the TB Photocatalyst

In this research, at the first step, the photocatalyst was synthesized by the co-precipitation protocol. The purpose of the photocatalyst manufacturing is to enhance the accessible surface for the TiO_2 and diminish the extra cost so that the catalyst can be operated with a low-priced base. To reach this aim, compounds of 5 wt %, 10 wt %, and 20 wt % of TiO_2 and bentonite were formulated. The mixture was stirred gently for 20 hours and by adding ethanol, a gel-like mixture was created. The created precipitate was washed with distilled water and the prepared gel was dried for about 2 h in the oven. The dried precipitate was heated in the furnace for 4 h at 450 ° C. Finally, the precipitate was slowly cooled down and sifted and for better sample recognition and tracking of the experimental results the catalysts were named TB 5, TB 10, TB 20 [21, 22].

2. 3. UV Irradiation Experiments

A solution containing 50 mg/L of PNT and 0.2 g/L TB was prepared and circulated for 120 minutes in the reactor with irradiation. samples were taken at certain reaction intervals, centrifuged and then analyzed by a UV–vis spectrophotometer at 286 nm.The degradation of the studied wastewater by the process was estimated using the following Equation (1). According to photocatalytic processes, when TiO_2 excitation occurs, it creates an electron and electron hole. The hole next to the water causes the production of OH^0 radicals. These radicals are very active and carry PNT and convert it into minerals. Figure 1 illustrated the degradation mechanism.

$$Degradation\ of\ PNT(\%) = \frac{C_0 - C}{C_0} \times 100 \tag{1}$$

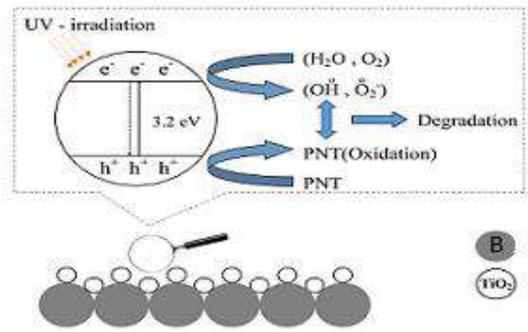


Figure 1. Schematic of the degradation mechanism

where C_0 and C are initial and current concentration of PNT, respectively.

3. RESULTS AND DISCUSSIONS

3. 1. Evaluation of Modified TB Photocatalyst Properties

3. 1. 1. TB Samples XRD Test

Figure 2 depicted the XRD diffraction pattern of the synthetic samples. The spectrum of the bentonite sample is shown in section (a). As illustrated in this figure, the sample consists of three main phases: Montmorlite, Quartz, and Feldspar. Section (b) fully complies with the formation of TiO_2 in the Anatase phase with the JCPDS card number of 1272-21-00 [23]. The bandwidths also confirm that the particles are all in a nano-sized scale. For the nanocomposite sample, as it is shown in sections (c-e), the Anatase phase peak intensities soar with increasing TiO_2 percentage from 5 wt% to 20 wt%. This intensity variation occurs specifically in the crystalline plates (101), (200) and (105) [24]. The simultaneous presence of TiO_2 and bentonite diffraction verifies the favorable synthesis of 5 wt%, 10 wt% and 20 wt% nanocomposites. However, in some angles, there is a high overlap between the two distinct components.

3. 1. 2. TB Samples FTIR Test

As Figure 3 illustrated, the (a) spectrum demonstrates the TiO_2 sample. The peak of Ti-O vibration is shown in the range of 400-500 cm^{-1}. The peaks at 3350 cm^{-1} and 1640 cm^{-1} are related to O-H tensile and bending vibrations, respectively [25].

Figure 3(b) spectrum is related to the bentonite sample. The bands observed at 460 cm^{-1} and 540 cm^{-1} are related to the bending vibrations of Si-O-Si and Al-O-Si. The peak observed at 626 cm^{-1} is attributed to off-plane vibrations of Al-O and Si-O bonds. The peak observed at 915 cm^{-1} is related to the tensile vibration of Al – O– (OH) –Al. The strong peak in the 1040 cm^{-1}

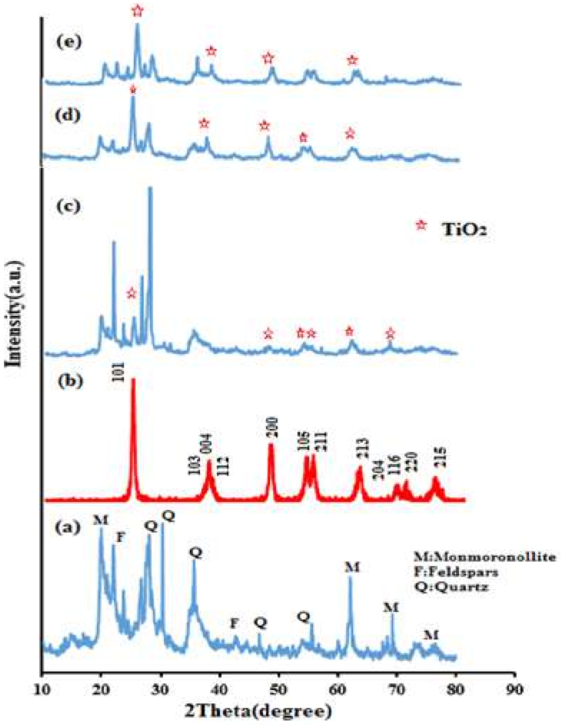


Figure 2. XRD patterns of (a) Bentonite , (b) Pure TiO_2, (c) TB 5, (d) TB 10 and (e) TB 20

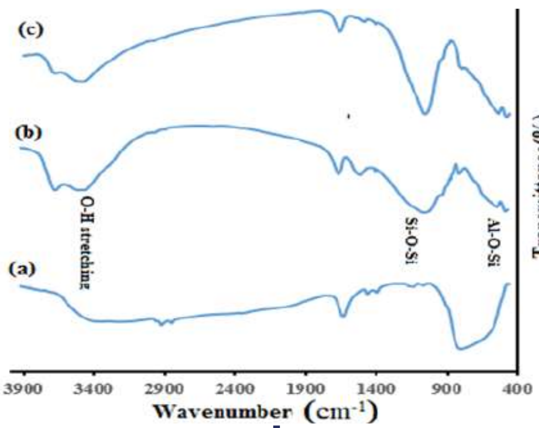


Figure 3. FT-IR spectra of (a) Pure TiO_2, (b) Pure Bentonite, (c) TB 20

region belongs to the Si-O-Si group of tetrahedral sheets. The peak observed at 1645 cm^{-1} is related to the bending vibration of the hydroxyl group of water molecules remaining in the bentonite matrix. The broadband at the 3420 cm^{-1} is attributed to the O-H tensile vibrations of the silanols (Si-OH) as well as the absorbed water [26].

The spectrum in Figure 3(c) highlights the composite sample of the TB 20 catalyst. The Ti-O peak is shown in the range of 400-500 cm^{-1} and the other peaks of the bentonite sample are fully preserved. However, in some cases, small displacements are observed at the peak location, which confirms the successful formation of the composite. These results are in full agreement with XRD and EDX analyses.

3. 1. 3. TB Sample EDX and SEM Tests Figure 4 depicts the SEM test for TB 5, TB 10, and TB 20 samples. This figure well indicates the presence of TiO_2 particles on the bentonite particles. The catalyst morphology is fixed in all the proportion of the compounds; likewise, the dispersion of TiO_2 particles is uniform on the surface. Figures illustrate that as the amount of TiO_2 enhanced, the number of particles on the surface escalated.

The EDX test was used for determination of different elements in various materials [27]. On the other hand, EDX tests were performed for the samples, and the results are shown in Figure 5.

The peaks imply the identification of major catalyst materials such as Si, C, Br, Ti and O. The results confirm the existence of Ti elements in the fabricated samples (peak 0.2 and 4.5 KeV) [28]. The composition percentage of the major elements is listed in Table 1.

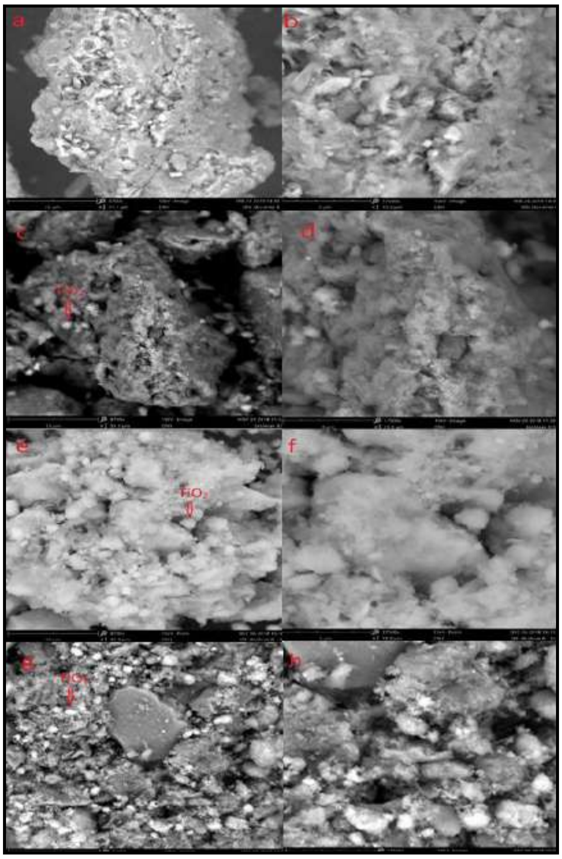


Figure 4. SEM of (a) Pure Bentonite (10 µm), (b) Pure Bentonite (5 µm), (c) TB 5 (10 µm), (d) TB 5 (5 µm), (e) TB 10 (10 µm), (f) TB 10 (5 µm), (g)TB 20 (10 µm), (h) TB 20 (5 µm)

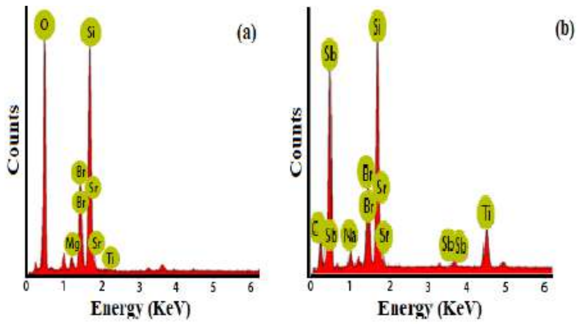


Figure 5. EDX patterns of (a) Bentonite, (b) TB 20

TABLE 1. Elemental chemical analysis of the synthesized samples

Element Symbol	Bentonite	TB 20	TB 10	TB 5
Si	22.4	14.9	16.9	21.4
O	49.8	43.5	43.7	49.4
Sr	6.2	3.9	4.2	5.9
Al	5.2	12.8	12.5	8.0
Ti	0.0	8.9	8.2	6.3
Na	2.4	2.0	2.0	2.3
Mg	1.6	1.3	1.3	1.6
Fe	2.4	1.1	1.0	3.4
Sb	2.3	1.8	1.8	1.7

3. 2. Evaluation of the Modified TB Photocatalyst on PNT Removal Performance After fulfilling the characteristics tests for evaluation of the synthesis method and well-creation of the hybrid photocatalyst, its performance was measured on the degradation of PNT. Experiments were carried out in the reactor setup. The prepared solution contained 50 mg/L of PNT, and the pH of the solution was kept at 5 (normal solution pH). These values were consistent by considering the previous results in all experiments. Also, the catalyst content was 0.2 g/L [20].

It should be noted that the intended values are optimal so that the rate of degradation depends on the initial concentration of paranitrotoluene. It changes with increasing initial concentration of the pollutant. The reaction between the hydroxyl radicals produced by the active sites on the surface of titanium dioxide and the paranitrotoluene molecule occurs from the solution.

When the initial contaminant content is high, the number of these active sites available by the paranitrotoluene molecules decreases as their absorption onto the surface of titanium dioxide decreases. On the other hand, there is an increase in the rate of transfer of paranitrotoluene molecules. When the initial concentration of paranitrotoluene is low, although there are more active sites and more hydroxyl radicals, the rate of contaminant transport to the catalyst surface is low, so it can be seen that the contaminant concentration has positive and negative effects [29].

pH has an important effect, and at low and high pH values, the rate of degradation is variable. The best pH for degradation is near the zero point of titanium dioxide, and this concept illustrates the effect of pH on the amount of degradation based on its effect on titanium dioxide particles. The zero point for titanium dioxide is in the pH range between 5.6 and 6.4 [30,31]. It appears that the maximum value of PNT destruction happens at the Neutral pH which may be originated from the superior production of total radicals at this pH. Truly, the PNT removal efficiency was improved due to keeping higher levels of radicals which are employed to decompose the structure of a refractory organic compound such as PNT.

The results showed that the rate of degradation increased with increasing concentration of titanium dioxide and then decreased. Excessive catalyst overload may cause solution obstruction and light penetration. Therefore, the catalyst increase has a positive effect until the turbidity effect is dominated, but afterwards it has a completely undesirable effect on the degradation process [32-34].

As it is shown in Figure 6, the degradation rate enhanced with increasing TiO_2 content and reaches its highest value of 59% for TB 20.

It is noteworthy that the degradation was very slow in the first half-hour due to the opacity of the solution, which diminished the illumination and consequently shortened photocatalytic activity. The highest catalytic activity occurs in the early hours and then the catalytic activity decreases and the reaction does not sensible change in the final hours. By the way, in the experiments that utilized pure TiO_2, the catalytic activity was low in the early hours and thereafter, the catalytic activity increased, with a maximum degradation rate of 64%.

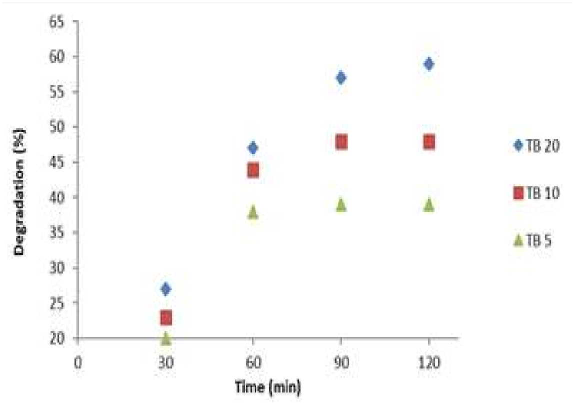


Figure 6. Photo catalytic degradation of PNT

4. CONCLUSION

In this study, the TB catalyst was synthesized by precipitation method, and its catalytic activity on PNT decomposition was explored. The catalyst was made of 5 wt%, 10 wt% and 20 wt% TiO_2 Furthermore, the characterizations of the nano catalyst were evaluated by the XRD, FT-IR and SEM techniques. Observations of this work reveal that modification of TiO_2 nano catalyst with Bentonite has a prominent impact on the degradation of PNT in a suspension reactor for the PNT values of 50 mg/L and TB of 0.2 g/L. The residence time of two hours was considered and it was observed that in the early hours of the reaction, the highest catalytic activity attains to develop and then the catalytic activity reduces and finally fixed at the equilibrium value. The results confirm that with increasing TiO_2, degradation rate increased and the highest value for TB 20 catalysts was 59% (COD = 53%). Therefore, it can be said that although the usage of bentonite as a base does not have much impact on enhancing the final efficiency, it can have a favorable role in terms of economic cost and reduction of reaction time.

5. ACKNOWLEDGEMENT

The authors wish to thank the Islamic Azad University, Arak Branch for the financial and laboratory equipment support.

6. REFERENCES

1. Mohamadiyan, J., Shams-Khoramabadi, G., Mussavi, SA., Kamarehie, B., Dadban Shahamat, Y., Godini, H., "Aniline degradation using advanced oxidation process by UV/Peroxy disulfate from aqueous solution", ***International Journal of Engineering, Transactions B: Applications***, Vol. 30, No. 5 (2017), 684-690, doi: 10.5829/idosi.ije.2017.30.05b.08
2. Mohadesi, M. and Shokri, A., "Treatment of oil refinery wastewater by photo-Fenton process using Box–Behnken design method: kinetic study and energy consumption", ***International Journal of Environmental Science and Technology***, Vol. 16, No. 11, (2019), 7349-7356.,
3. Taghavi, K., Naghipour, D., Mohagheghian, A and Moslemzadeh, M., "Photochemical Degradation of 2,4-dichlorophenol in Aqueous Solutions by Fe^{2+}/Peroxydisulfate/UV Process", ***International Journal of Engineering, Transactions A: Basics***, Vol. 30, No. 1, (2017), 15-22, doi: 10.5829/idosi.ije.2017.30.01a.03
4. Ludwichk, R., Helferich, O.K., Kist, C.P., Lopes, A.C., Cavasotto, T., Silva, D.C. and Barreto-Rodrigues, M., "Characterization and Photocatalytic Treatability of Red Water from Brazilian TNT Industry", ***Journal of Hazardous Materials***, Vol. 293, (2015), 81-86, doi:10.1016/j.jhazmat.2015.03.017.
5. Shokri, A., "Employing Sono-Fenton Process for Degradation of 2-Nitrophenol in Aqueous Environment Using Box–Behnken Design Method and Kinetic Study", ***Russian Journal of Physical Chemistry A***, Vol. 93, No. 2, (2019), 243-249, doi:10.1134/S003602441902002X.
6. Hosseini, J. and Shokri, A., "Employing UV/H_2O_2 Process for Degradation of 2, 4-Diaminotoluene in Synthetic Wastewater", ***Archives of Hygiene Sciences***, Vol. 6, No. 2, (2017), 121-127,
7. Shokri, A. and Mahanpoor, K., "Removal of Ortho-Toluidine from Industrial Wastewater by UV/TiO_2 Process", ***Journal of Chemical Health Risks***, Vol. 6, No. 3, (2016), 213-223, doi:10.22034/jchr.2016.544149.
8. Ayoub, K., van Hullebusch, E.D., Cassir, M. and Bermond, A., "Application of Advanced Oxidation Processes for TNT Removal: A Review", ***Journal of Hazardous Materials***, Vol. 178, No. 1-3, (2010), 10-28, doi:10.1016/j.jhazmat.2010.02.042.
9. Huang, N., Wang, T., Wang, W.-L., Wu, Q.-Y., Li, A. and Hu, H.-Y., "UV/Chlorine as an Advanced Oxidation Process for the Degradation of Benzalkonium Chloride: Synergistic Effect, Transformation Products and Toxicity Evaluation", ***Water Research***, Vol. 114, (2017), 246-253, doi:10.1016/j.watres.2017.02.015.
10. Lee, S.-Y. and Park, S.-J., "TiO_2 Photocatalyst for Water Treatment Applications", ***Journal of Industrial and Engineering Chemistry***, Vol. 19, No. 6, (2013), 1761-1769, doi:10.1016/j.jiec.2013.07.012.
11. Qi, W., Zhang, F., An, X., Liu, H. and Qu, J., "Oxygen Vacancy Modulation of {010}-dominated TiO_2 for Enhanced Photodegradation of Sulfamethoxazole", ***Catalysis Communications***, Vol. 118, (2019), 35-38, doi:10.1016/j.catcom.2018.09.014.
12. Trckova, M., Matlova, L. and Dvorska, L., "Kaolin, Bentonite, and Zeolites as Feed Supplements for Animals: Health Advantages and Risks: A Review", ***Veterinarni Medicina-UZPI (Czech Republic)***, (2004),
13. Nasirian, M., Bustillo-Lecompte, C.F. and Mehrvar, M., "Photocatalytic Efficiency of Fe_2O_3/TiO_2 for the Degradation of Typical Dyes in Textile Industries: Effects of Calcination Temperature and UV-assisted Thermal Synthesis", ***Journal of Environmental Management***, Vol. 196, (2017), 487-498, doi:10.1016/j.jenvman.2017.03.030.
14. Setthaya, N., Chindaprasirt, P., Yin, S. and Pimraksa, K., "TiO_2-zeolite Photocatalysts Made of Metakaolin and Rice Husk Ash for Removal of Methylene Blue Dye", ***Powder Technology***, Vol. 313, (2017), 417-426, doi:10.1016/j.powtec.2017.01.014.
15. Hoseini, S.N., Pirzaman, A.K., Aroon, M.A. and Pirbazari, A.E., "Photocatalytic Degradation of 2,4-dichlorophenol by Co-doped TiO_2 (Co/TiO_2) Nanoparticles and Co/TiO_2 Containing Mixed Matrix Membranes", ***Journal of Water Process Engineering***, Vol. 17, (2017), 124-134, doi:10.1016/j.jwpe.2017.02.015.
16. Gharibshahian, E., "The Effect of Polyvinyl Alcohol Concentration on the Growth Kinetics of $KTiOPO_4$ Nanoparticles Synthesized by the Co-precipitation Method", ***HighTech and Innovation Journal***, Vol. 1, No. 4, (2020), 187-193, doi:10.28991/HIJ-2020-01-04-06.
17. Theingi, M., Tun, K.T. and Aung, N.N., "Preparation, Characterization and Optical Property of $LaFeO_3$ Nanoparticles via Sol-Gel Combustion Method", ***SciMedicine Journal***, Vol. 1, No. 3, (2019), 151-157, doi:10.28991/SciMedJ-2019-0103-5.
18. Aby, H., Kshirsagar, A. and Khanna, P., "Plasmon Mediated Photocatalysis by Solar Active Ag/ZnO Nanostructures: Degradation of Organic Pollutants in Aqueous Conditions", ***J Mater Sci Nanotechnol***, Vol. 4, No. 1, (2016), 103,
19. Rosales, B.S., Moreira del Rio, J., Guayaquil, J.F. and de Lasa, H., "Photodegradation Efficiencies in a Photo-CREC Water-II Reactor Using Several TiO_2 Based Catalysts", ***International Journal of Chemical Reactor Engineering***, Vol. 14, No. 3, (2016), 685-701, doi:10.1515/ijcre-2016-0024.

20. Rostami, M., Mazaheri, H., Hassani Joshaghani, A. and Shokri, A., "Using Experimental Design to Optimize the Photo-degradation of P-Nitro Toluene by Nano-TiO_2 in Synthetic Wastewater", ***International Journal of Engineering, Transactions B: Applications,*** Vol. 32, No. 8, (2019), 1074-1081, doi:10.5829/ije.2019.32.08b.03.
21. Mahlambi, M.M., Mishra, A.K., Mishra, S.B., Raichur, A.M., Mamba, B.B. and Krause, R.W., "Layer-by-Layer Self-Assembled Metal-Ion-(Ag-, Co-, Ni-, and Pd-) Doped TiO_2 Nanoparticles: Synthesis, Characterisation, and Visible Light Degradation of Rhodamine B", ***Journal of Nanomaterials***, (2012), doi:10.1155/2012/302046.
22. Daghrir, R., Drogui, P. and Robert, D., "Modified TiO_2 For Environmental Photocatalytic Applications: A Review", ***Industrial & Engineering Chemistry Research***, Vol. 52, No. 10, (2013), 3581-3599, doi:10.1021/ie303468t.
23. Wei, X., Zhu, G., Fang, J. and Chen, J., "Synthesis, Characterization, and Photocatalysis of Well-Dispersible Phase-Pure Anatase TiO_2 Nanoparticles", ***International Journal of Photoenergy***, Vol. 2013, (2013), doi:10.1155/2013/726872.
24. Malamis, S. and Katsou, E., "A Review on Zinc and Nickel Adsorption on Natural and Modified Zeolite, Bentonite and Vermiculite: Examination of Process Parameters, Kinetics and Isotherms", ***Journal of Hazardous Materials***, Vol. 252-253, (2013), 428-461, doi:10.1016/j.jhazmat.2013.03.024.
25. Kumar, P.M., Badrinarayanan, S. and Sastry, M., "Nanocrystalline TiO_2 Studied by Optical, FTIR and X-ray Photoelectron Spectroscopy: Correlation to Presence of Surface States", ***Thin Solid Films***, Vol. 358, No. 1, (2000), 122-130, doi:10.1016/S0040-6090(99)00722-1.
26. Paluszkiewicz, C., Holtzer, M. and Bobrowski, A., "FTIR Analysis of Bentonite in Moulding Sands", ***Journal of Molecular Structure***, Vol. 880, No. 1, (2008), 109-114, doi:10.1016/j.molstruc.2008.01.028.
27. Shokri, A., "An Investigation of Corrosion and Sedimentation in the Air Cooler Tubes of Benzene Drying Column in Linear Alkyl Benzene Production Plant", ***Chemical Papers***, Vol. 73, No. 9, (2019), 2265-2274,
28. Li, W.-C., Comotti, M. and Schüth, F., "Highly Reproducible Syntheses of Active Au/TiO_2 Catalysts for Co Oxidation by Deposition–Precipitation or Impregnation", ***Journal of Catalysis***, Vol. 237, No. 1, (2006), 190-196, doi:10.1016/j.jcat.2005.11.006.
29. Shokri, A. and Mahanpoor, K., "Degradation of Ortho-Toluidine from Aqueous Solution by the TiO_2/O_3 Process", ***International Journal of Industrial Chemistry***, Vol. 8, No. 1, (2017), 101-108, doi:10.1007/s40090-016-0110-z.
30. Shokri, A., Moradi, H., Abdouss, M. and Nasernejad, B., "Employing UV/Periodate Process for Degradation of p-chloronitrobenzene in Aqueous Environment", ***Desalination and Water Treatment***, Vol. 205, (2020), 264-274, doi:10.5004/dwt.2020.26384.
31. Bayarri, B., Gimenez, J., Curco, D. and Esplugas, S., "Photocatalytic Degradation of 2, 4-dichlorophenol by TiO_2/UV: Kinetics, Actinometries and Models", ***Catalysis Today***, Vol. 101, No. 3-4, (2005), 227-236, doi:10.1016/j.cattod.2005.03.019.
32. Shokri, A. and Joshagani, A.H., "Using Microwave Along with TiO_2 for Degradation of 4-chloro-2-nitrophenol in Aqueous Environment", ***Russian Journal of Applied Chemistry***, Vol. 89, No. 12, (2016), 1985-1990, doi:10.1134/S1070427216120090.
33. Beridze, M., Kalandia, A., Japaridze, I., Vanidze, M., Varshanidze, N., Turmanidze, N., Dolidze, K., Diasamidze, I. and Jakeli, E., "Phytochemical Study of Endemic Species Helleborus Caucasicus and Helleborus Abchasicus", ***HighTech and Innovation Journal***, Vol. 1, No. 1, (2020), 28-32, doi:10.28991/HIJ-2020-01-01-04.
34. Manikandan, G., Yuvashree, M., Sangeetha, A., Bhuvana, K. and Nayak, S.K., "Liver Tissue Regeneration Using Nano Silver Impregnated Sodium Alginate/PVA Composite Nanofibres", ***SciMedicine Journal***, Vol. 2, No. 1, (2020), 16-21, doi:10.28991/SciMedJ-2020-0201-3.

Persian Abstract

چکیده

کشور ایران در دهه های اخیر با مشکل جدی کمبود آب مواجه است. از آنجایی که صنعت، یکی از بخش های پر مصرف آب به شمار می رود، تصفیه پساب های صنعتی امری ضروری به نظر می رسد. از طرفی مشتقات نیتروتولوئن کاربرد زیادی در صنایع، بویژه صنعت نظامی دارد که خود نیز سهم گسترده ای در آلودگی پساب های صنعتی دارد. این ترکیبات به شدت برای موجودات زنده خطرناک است و میتواند آثار جبران ناپذیری را از خود به جا گذارد از این رو حذف آنها در پساب صنعتی امری ضروری می باشد . فرآیندهای فتوکاتالیستی از زیر مجموعه فرآیندهای اکسایش پیشرفته یکی از روش های نوین در تصفیه پساب های صنعتی به شمار می رود. یکی از پرکاربرد ترین فتوکاتالیست ها در این فرآیند دی اکسید تیتانیوم است. در این تحقیق اصلاح کاتالیزور دی اکسید تیتانیوم با هدف صرفه اقتصادی و با حفظ شرایط تخریب مورد بررسی قرار گرفت. به این منظور ترکیب بنتونیت و دی اکسید تیتانیوم (TB) به روش هم رسوبی ساخته وفعالیت کاتالیزوری آن بر روی تخریب پارانیتروتولوئن بررسی شد. کاتالیزور اصلاح شده با نسبت 5، 10 و 20 درصد از دی اکسید تیتانیوم ساخته و خواص آن توسط خصوصیات توسط میکروسکوپ الکترونی روبشی (SEM) ، پراش اشعه ایکس (XRD) ، طیف سنجی مادون قرمز تبدیل فوریه (FT-IR) و طیف سنجی پراش انرژی پرتو ایکس (EDX) تعیین گردید. نتایج نشان می دهد که با افزایش مقدار دی اکسید تیتانیوم، میزان تخریب افزایش یافته و بیشترین مقدار آن برای کاتالیزور TB 20 به میزان 59 درصد می باشد. این در حالی است که ، تحت همان شرایط ، برای TiO2 خالص ، میزان تخریب 64٪ است اما با این تفاوت که زمان واکنش و میزان مصرف TiO_2 در این حالت بیشتر است .سرانجام ، به منظور تأیید بیشتر میزان تخریب ، آزمایش اکسیژن مورد نیاز شیمیایی (COD) بر روی نمونه TA 20 انجام شد. نتایج نشان داد که حدود 53٪ از آلاینده اولیه به مواد معدنی تبدیل شده است.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 763-768

International Journal of Engineering

Journal Homepage: www.ije.ir

Risk Assessment of Gasoline Storage Unit of National Iranian Oil Product Distribution Company using PHAST Software

H. Esfandian*[a], M. Goodarzian Urimi[b], A. Shokoohi Rad[b]

[a] *Faculty of Engineering Technologies, Amol University of Special Modern Technologies, Amol, Iran*
[b] *Department of Chemical Engineering, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran*

PAPER INFO

Paper history:
Received 05 October 2020
Received in revised form 19 October 2020
Accepted 05 November 2020

Keywords:
Gasoline Reservoirs
Consequences Modeling
Risk Assessment
F-N Curve
PHAST

ABSTRACT

The present study evaluates the risk of the gasoline tank of the National Iranian Oil Product Distribution Company (NIOPDC) in Sari region using process hazard analysis software tool (PHAST) and according to the environmental and process data of the unit. The consequences of different scenarios such as small and medium leakage, constant release rate and complete rupture were modeled and then the range of each one was obtained according to the intensity of radiation or pressure wave and the safe distances of each was determined. Due to the consequences of the explosion, the worst results were related to the weather conditions of 2/3 F for 4700, 2400, and 2300 meters, respectively. Also, based on eruptive and sudden fire data, the intensity of radiation which corresponds to the immediate death or destruction of equipment was seen in climatic conditions of (2/3 F and 4/1 D), at intervals of 180 and 160 meters distances, respectively. In these two weather conditions flammability intervals were 10520 and 450 meters. Then, by combining the severity of these accidents with the distribution of the population and the probability of their occurrence, the level of risk for these storages was determined.

***doi**: 10.5829/ije.2021.34.04a.02*

1. INTRODUCTION

Since the desire for safety and security is an integral part of the nature of human beings, having a life free of danger has always been the desire and objective of all people. The development of various industries enhanced the well-being of human beings, but they created new potential dangers as well. Furthermore, the process facilities and equipment that are used for productivity, profitability and wealth creation in various industries, have a great potential for causing harm to people, property and the environment. Risk analysis and assessment is one of the most important tools to maintain and improve the level of safety in the society and especially in industry [1-3]. For this purpose, in existing or under design industrial units, risk assessment is performed for the hazards that may occur due to human error or equipment failure. Storage tanks in refineries and petrochemical units contain high amounts of hazardous and sometimes flammable substances [4, 5]. These tanks are fragile and easily damaged by a slight increase in pressure or vacuum, therefore, they are more prone to accidents than other equipment. As a result, a small accident can cause millions of dollars in property damage and stop the production process, and it also may result in loss of life. Given the events that have taken place in the case of storage tanks in recent years, the importance of assessing and evaluating the risk of storage tanks in refinery units is completely clear. Based on the available resources, 242 accidents have occurred in the chemical tanks of industrial facilities in worldwide, during the last 40 years. Of the 242 accidents, 114 cases were occurred in North America, 72 cases in Asia and 38 cases in the Europe. The highest number of accidents, 116 cases (47.8%), were occurred in the oil refineries, 64 cases (26.4%) in oil terminals and loading platforms and about 25.7% in petrochemical units. The results of this study show that 74% of the mentioned accidents in the oil industry are related to the oil storage and loading terminals. The highest number of causes of accidents are

*Corresponding Author Institutional Email: *H.Esfandian@ausmt.ac.ir* (H. Esfandian)

Please cite this article as: H. Esfandian, M. Goodarzian Urimi, A. Shokoohi Rad, Risk Assessment of Gasoline Storage Unit of National Iranian Oil Product Distribution Company using PHAST Software, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 763-768

fire and explosion in these tanks that were 85% of all the cases. About 33% of accidents were related to lightning and 30% of them were due to human error, including operational and repair errors. Other causes of accidents include equipment failure, vandalism, rupture, leakage and rupture of lines, static electricity and open fire around tanks. Gasoline storage tanks, as one of the most important industrial facilities, are always exposed to the risk of fire and explosion, therefore, risk assessment of these tanks is an effective step to prevent accidents or reduce their severity [4, 6]. Unfortunately, in our country it is normally ignored. Perhaps among the important reasons is the lack of sufficient familiarity with the basic principles of quantitative and qualitative risk assessment. One of the most important and common accidents in industrial units, public buildings and crowded places that endanger human lives and damage equipment and tools, is the phenomenon of fire and explosion [7]. Some of the incidents include the gasoline fire in Cubatao, Brazil, which killed more than 500 people; the Bhopal disaster in India, which killed at least 2,500 people; the reservoir explosion in Mexico, which injured 4,200 persons and killed 452 individuals. The explosion of flammable material in Mexico's sewage network, which was estimated to cost more than $ 7 billion. As a result, these accidents have caused human, equipment and environmental catastrophes, and sometimes these damages are heavy and irreparable. Such an events are causing more concern and people are thinking about the consequences and effects of process accidents in industrial activitie [7]. A study conducted in the field of energy in recent years [8-11] indicated that accidents caused by explosions and fires in the oil industry accounted for 25% of the total economic losses. It is in the second place, after nuclear energy. The high human, economic and environmental damage reported from accidents can be a very good reason to show the need to observe safety principles, not only in chemical units but in all fields [12]. In the present study the generalities of risk assessment and its application, including the steps of identifying hazards in a process, modeling potential hazards such as fire, explosion or consequences related to toxicity of materials, risk assessment and its extent will be described. The risk assessment will be done specifically for the gasoline storage unit of NIOPDC in the Sari region. PHAST specialized software is used to model the consequences of accidents in this unit. Finally, after performing the calculations of probable risk of accidents in this unit and comparing it with valid criteria, appropriate suggestions will be provided to reduce the risk of relevant accidents.

2. MATERIALS AND METHODS

This study uses PHAST software. PHAST software is one of the most popular and useful accident modeling software related to the release of toxic substances, fire and explosion [13-16]. This software is provided by DNV Company and is well known for its industrial and public safety hazards. This model is one of the best models presented for the distribution of materials in the environment. This model covers a wide range of pure materials lighter or heavier than air and it is able to model a mixture of materials. It also includes sudden, permanent release and evaporation from the surface of the ponds. Release height and average ground surface roughness are considered in this model. In this article, the risk assessment of the gas tank of the NIOPDC in Sari region is examined. These tanks are made of carbon steel, and they are in the cylinders form with a height and diameter of 12.84 and 34.178 meters, respectively, and the total volume of each of them is 10,000 cubic meters. There are 8 similar tanks next to each other in the studied unit.

2. 1. Modeling the Consequences of Scenarios

Scenario definition is one of the first steps in risk assessment that predicts possible events. In the present study, the scenarios are defined as follows: Due to wear and defects that may occur in the system, there is a possibility of kerosene leakage from various parts of the source, such as leaks from flanges, washers, and connecting pipes, sudden puncture and bursting of the tank. Depending on the kerosene conditions, the consequences of the blast wave and the intensity of the current radiation from the reservoir events are investigated. The environmental consequences of the release of kerosene from storage tanks into the environment and the destruction of the environment are among the chronic hazards that affect human life and other organisms in the long run.

The different scenarios that are examined in this study are:

- Leakage with small diameters (10 mm)
- Complete rupture of the tank and its sudden discharge
- Release of the entire materials in the tank at a constant release rate in 10 minutes

Considering that one of the stages of modeling the release of materials in the environment is the climatic characteristics of the environment, Table 1 indicates the seasonal average conditions (hot-cold) climate for the Sari region and the environmental conditions of the region. Therefore, for modeling, the outcome of each of these conditions is considered as modeling conditions. The results of the outcome modeling were obtained according to the studied unit conditions and environmental characteristics and according to the equations expressed in the third chapter.

3. RESULTS AND DISCUSSION

3. 1. The First Scenario - Small Leak

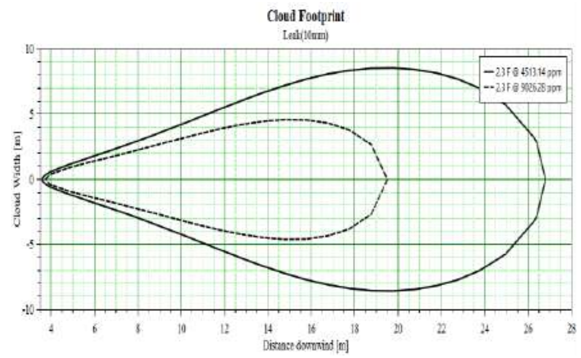
The modeling results in climatic conditions (2.3 F and 4.1 D)

TABLE 1. Climatic and Environmental Conditions of the Region

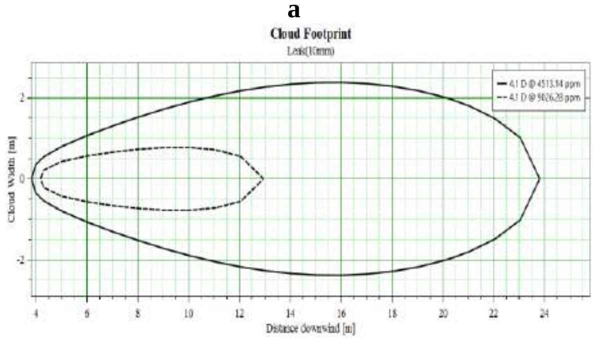
	Hot	Cold
Average Temperature ℃	13	25
Wind Speed (m/s)	2.3	4.1
Sustainability Class	F	D
Humidity	0.9	0.75
Radiation (kW/m^2)	0.4	0.7
Surface Roughness (mm)	183	183

for the leakage scenario indicate a sudden fire, eruptive fire, and explosion. Figure 1 (a and b) show the amplitude of the cloud concentration created on the land surface from the top view in terms of distance from the source and the width of the cloud created in 2.3 F and 1.4 D weather conditions. Flammability concentrations of 9026 and 4513 ppm for climates 2.3 F and 1.4 D are advanced to a distance of 19.5 - 27 m and 13- 24 m from the reservoir, respectively. The cloud width created for the lower limit of flammability in two weather conditions of 2.3 F and 1.4 D is 18 and 5 meters, respectively. According to the results, in 2.3 F weather conditions, due to the lower wind speed and more stable atmospheric conditions, the cloud is created to spread at a greater distance. Figure 2 (a and b) indicate the enclosed area to the concentration of fire. The half of this concentration is for sudden fire in weather conditions of 2.3 F and 4.1 D. In these figures the horizontal axis is the distance from the release source in the wind direction; the vertical axis is the distance from the release source in the direction perpendicular to the wind and parallel to the earth's surface. These diagrams are to determine the geographical location of the range of radiation caused by a sudden fire. According to the results, in weather conditions 2.3 F a zone with a radius of 27 meters, which is marked with a line, has a concentration of more than half of the lower flammability limit. On the other hand, a zone with a radius of 20 meters has a concentration more than the flammability limit. While these distances are 4.1, the radius is 24 m and 14 m for weather conditions of 4.1 D. In other words, as the wind speed increases, the scattered cloud concentration dilutes faster, reducing the ignition concentration parameter limit.

Figure 3 (a and b) show the radiation from the eruptive fire based on the distance in the wind direction for a leak with a diameter of 10 mm. In these figures, the different levels of thermal radiation specified by the model for eruptive fire are shown as limits in terms of distance. In these diagrams, after determining the desired radiation intensities, the geographical areas in which the minimum radiation intensity is equal to the desired values are determined. As can be seen, for climatic conditions 2.3 F and radiation 4, 12.5, and 37.5 kW/m it continues

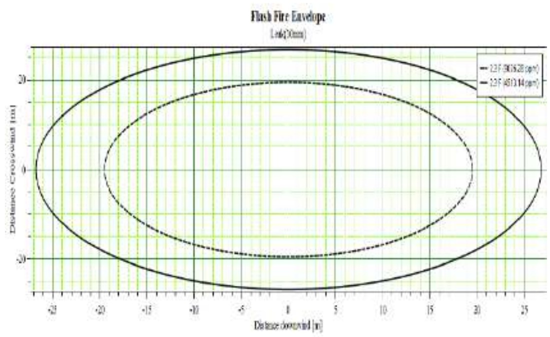


a

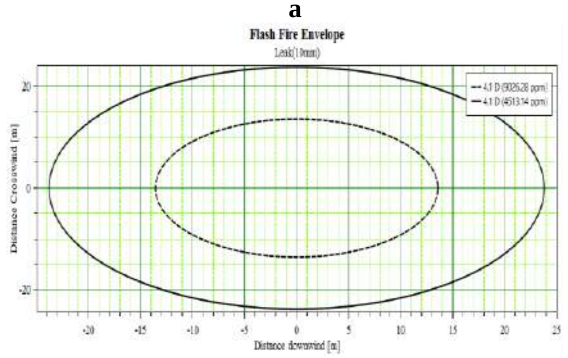


b

Figure 1. Vapor cloud emission range for the first scenario in **a**- 2.3 F, **b**- 4.1 D.

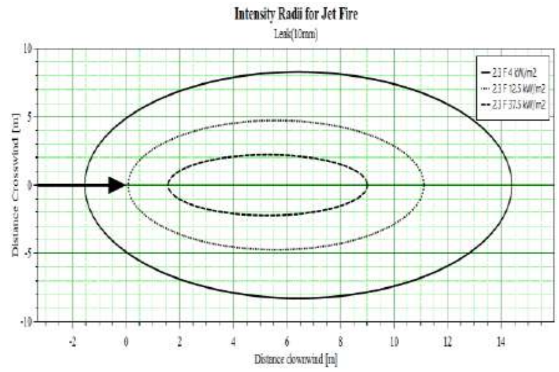


a

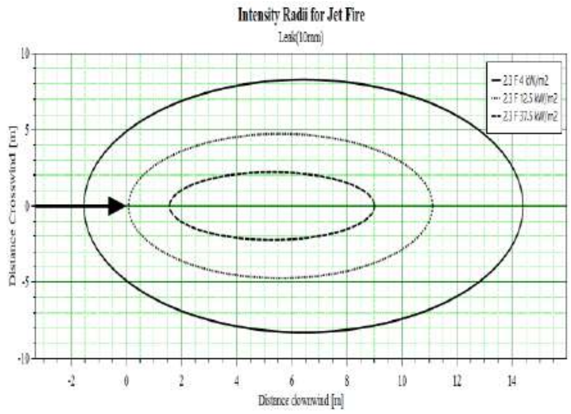


b

Figure 2. Sudden fire emission range for the first scenario in **a**- 2.3 F, **b**- 4.1 D.



a



b

Figure 3. Eruptive fire emission range for the first scenario in **a**-2.3 F, **b**-4.1D

up to radius 14.5, 11, and 9 meters, respectively, and for climatic conditions 4.1 D it continues up to a radius of 13.5, 10, and 8 meters, respectively.

Figure 4 indicates that three different levels of vulnerability for fire eruptive fire in 2.3 F and 4.1 D climates. According to the data obtained, people or equipment that are approximately 8.5 meters radius from the release source is going to be destroyed due to radiation from the eruptive fire. The vulnerability of 1% and 10% in this scenario occurs at intervals of approximately 11.5 meters and 10 meters for equipment and staff. Figure 5 (a and b) shows the ranges related to the increase of pressure created by the explosion in the direction of the wind and three different values of 0.0207, 0.1379, and 0.2068 bar at distances of 60, 28, and 26 m from the reservoir for climate conditions of 2.3 F, and 47, 25, and 24 m for the climate condition of 4.1 D. Explosion wave values are determined to indicate the areas affected by them and to determine safe distances in the model and based on the sources and their consequences (window breakage, partial and complete destruction of the building, death, and so on).

3. 2. The Second Scenario-Complete Rupture

The results of complete rupture modeling are reported as

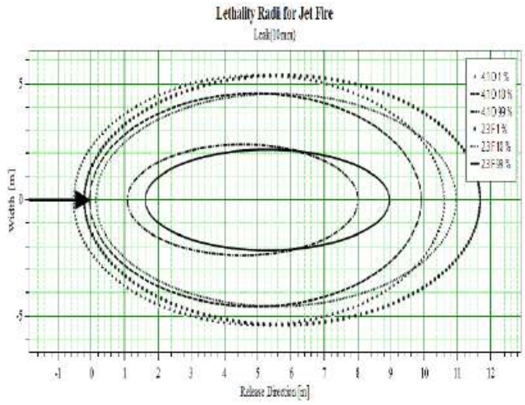
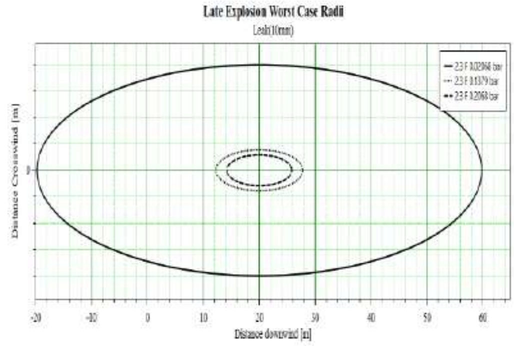
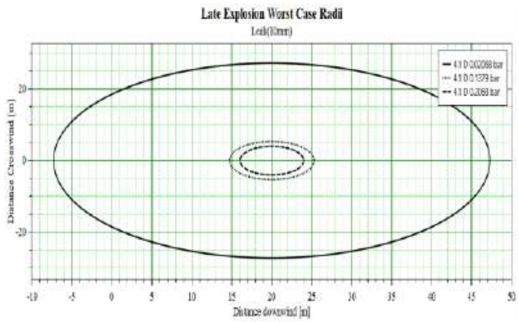


Figure 4. Eruptive fire emission range for the first scenario in 2.3 F and 4.1 D



a



b

Figure 5. Range of pressure increasing levels due to expulsion wave for the first Scenario in **a**-2.3 F, **b**-4.1D

the worst case scenario in simulating possible consequences in climatic conditions (2.3 F and 4.1 D). According to the results, the width of the formed cloud for fire concentration was obtained 800 and 670 m, and the cloud of vapor rises to approximate heights of 4 and 2 m in climatic conditions 2.3 F and 4.1 D respectively. In addition, in 2.3 F climates, the fire concentration and half of this concentration are confined to the distances of 1600 and 930 m, but due to the fact that in 4.1 D climates, the wind speed is higher and the atmosphere is more unstable, the digits are reduced to 1200 and 800 m, respectively. The ranges related to the increase in pressure created by the explosion in the wind direction at three different range of 0.0207, 0.1379, and 0.2068 bar was obtained at distances of 2700, 1300, and 1200m from the reservoir for water conditions 4.1 D, respectively. These values are 3600, 2000, and 1850 m for 2.3 F weather conditions, respectively.

3. 3. The Third Scenario-Continuous Leakage with a Constant Release Rate

In this section, the results of modeling continuous release at a constant rate for 10 minutes in different climatic conditions (2.3 F and 4.1 D) are expressed. In this scenario, the cloud formation of the ignition concentration expands to a distance of 1025 m for weather conditions 2.3 F and to a distance of 460 m for weather conditions 4.1 D. According to the results in climatic conditions 2.3 F, the range with a radius of 1900 meters has a concentration of more than half of the lower flammability limit, and the range with a radius of 1050 meters has a concentration more than the flammability limit, while these distances are for climate 4.1 D, radius 700 and 450 meters. The modeling results show that with increasing wind speed and decreasing atmospheric stability, the scattered cloud concentration dilutes faster; therefore, the radius corresponding to the fire concentration parameter is less in 4.1 D. In addition, results indicate that, the intensity of radiation, which corresponds to the immediate death or destruction of equipment, occurs in climatic conditions (2.3 F and 4.1 D), at intervals of 180 and 160 meters, respectively. So, at an approximate distance of more than 290 meters almost no radiation is observed from this consequence. The areas related to the increase in pressure created by the explosion of the constant release scenario in the wind direction and in three different values: 0.0207, 0.1379, and 0.2068 bar are at distances of 2400, 1000 and 920 meters from the reservoir, respectively, for weather conditions 4.1 D. These values are 4700, 2400, and 2300 meters for 2.3 F.

The results of risk assessment regarding the occurrence of different scenarios is indicated using F-N curves, by calculating the amount of Probity Function and the probability of mortality, assuming that the number of people in the affected area are scattered with an average population distribution coefficient of 0.001. Figure 6 indicates that the F-N curve for the studied scenarios. The vertical and horizontal axes in this figure show the reproducibility of scenarios and the number of casualties due to accidents over a period of one year. In fact, at this stage, a combination of repeatability and consequences of scenarios and mortality rates is used to determine the risk. In this diagram, the upper diagonal line shows the high-risk criterion and the lower diagonal line shows the low risk criterion. The area between the two is the average risk area. The broken line between the two diagonal lines represents the desired process in the unit under study.

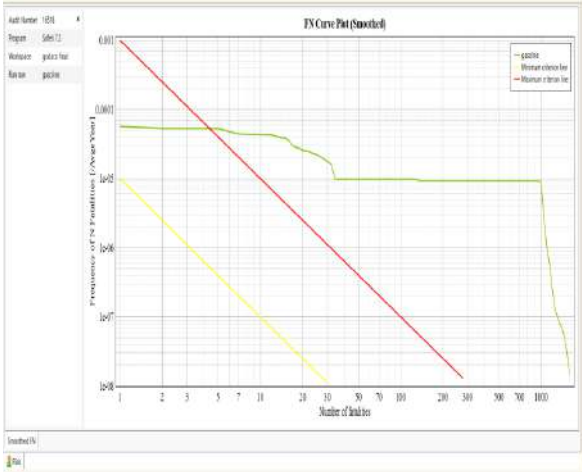


Figure 6. F-N Curve

4. CONCLUSION

The figure indicates that a high percentage of the process under study is in the high-risk area. To reduce the risk, two factors must be reduced, the consequence factor or its repeatability, or both. Many accidents occur due to corrosion and decay of connections and equipment. One of the strategies to reduce reproducibility includes increasing periodic inspections, thickness measurements, and monitoring. Given that the scenario of sudden discharge of tank contents and the consequence of explosion and fire is an instantaneous and unpredictable event, the time parameter has no effect on reducing the severity of casualties. In the contrary, the consequences of leakage events with a constant release rate or leakage of 10 mm is strongly dependent on the leakage time. Therefore, it is recommended that leak-sensitive sensors be installed near hazardous equipment to identify the leak as soon as possible and to eliminate it so that as a result, the amount of damage is minimized.

5. ACKNOWLEDGMENT

The authors would like to acknowledge Amol University of Special Modern Technologies for all the support.

6. REFERENCES

1. Wang, H., Khan, F., Ahmed, S. and Imtiaz, S., "Dynamic quantitative operational risk assessment of chemical processes", ***Chemical Engineering Science***, Vol. 142, (2016), 62-78. https://doi.org/10.1016/j.ces.2015.11.034

2. Wu, D. and Chen, Z., "Quantitative risk assessment of fire accidents of large-scale oil tanks triggered by lightning", ***Engineering Failure Analysis***, Vol. 63, (2016), 172-181. https://doi.org/10.1016/j.engfailanal.2015.11.029

3. Dadashzadeh, M., Kashkarov, S., Makarov, D. and Molkov, V., "Risk assessment methodology for onboard hydrogen storage", ***International Journal of Hydrogen Energy***, Vol. 43, No. 12, (2018), 6462-6475. https://doi.org/10.1016/j.ijhydene.2018.01.195
4. Fuentes-Bargues, J.L., González-Cruz, M., González-Gaya, C. and Baixauli-Pérez, M., "Risk analysis of a fuel storage terminal using hazop and fta", ***International Journal of Environmental Research Public Health***, Vol. 14, No. 7, (2017), 705. https://doi.org/10.3390/ijerph14070705
5. Ahmed, M.M., Ibrahim, M.M. and Al Ghabshawi, O.M., "Risk assessment for hydrocarbon fuel storage and handling facilities at gaili area, khartoum north-sudan", ***IRA-International Journal of Applied Sciences***, Vol. 10, No. 2, 18-26.
6. Swick, D., Jaques, A., Walker, J. and Estreicher, H., "Gasoline risk management: A compendium of regulations, standards, and industry practices", ***Regulatory Toxicology Pharmacology***, Vol. 70, No. 2, (2014), S80-S92. https://doi.org/10.1016/j.yrtph.2014.06.022
7. Badri, N., Nourai, F. and Rashtchian, D., "Quantitative risk assessment to site cng refueling stations", ***Chemical Engineering Transactions***, Vol. 19, (2010), 255. https://doi.org/10.3303/CET1019042
8. LaFleur, A.C., Muna, A.B. and Groth, K.M., "Application of quantitative risk assessment for performance-based permitting of hydrogen fueling stations", ***International Journal of Hydrogen Energy***, Vol. 42, No. 11, (2017), 7529-7535. https://doi.org/10.1016/j.ijhydene.2016.06.167
9. Khakzad, N., Dadashzadeh, M. and Reniers, G., "Quantitative assessment of wildfire risk in oil facilities", ***Journal of Environmental Management***, Vol. 223, (2018), 433-443. https://doi.org/10.1016/j.jenvman.2018.06.062
10. Xu, X., Wang, F., Huang, M., Bai, J. and Li, L., "Security quantitative risk analysis of ethylene horizontal tanks of a petrochemical company", ***Procedia Engineering***, Vol. 45, (2012), 489-495. https://doi.org/10.1016/j.proeng.2012.08.191
11. Nelson, T.P., "An examination of historical air pollutant emissions from us petroleum refineries", ***Environmental Progress Sustainable Energy***, Vol. 32, No. 2, (2013), 425-432. https://doi.org/10.1002/ep.11713
12. Khan, F.I. and Abbasi, S., "Modelling and control of the dispersion of hazardous heavy gases", ***Journal of Loss Prevention in the Process Industries***, Vol. 12, No. 3, (1999), 235-244. https://doi.org/10.1016/S0950-4230(98)00009-6
13. Taghehbaf, M.A., Givehchi, S., Ardestani, M. and Baghvand, A., *Modeling the consequences of potential accidents in one of the gasoline storage tanks at oil storage of yazd, in terms of explosion*. 2014, IJEIR.
14. Bouafia, A., Bougofa, M., Rouainia, M. and Medjram, M.S., "Safety risk analysis and accidents modeling of a major gasoline release in petrochemical plant", ***Journal of Failure Analysis Prevention***, Vol. 20, (2020), 358-369. https://doi.org/10.1007/s11668-020-00826-9
15. Zhou, J., You, Y., Bai, Z., Hu, Y., Zhang, J. and Zhang, N., "Health risk assessment of personal inhalation exposure to volatile organic compounds in tianjin, china", ***Science of the Total Environment***, Vol. 409, No. 3, (2011), 452-459. https://doi.org/10.1016/j.scitotenv.2010.10.022
16. Golbabai, F., Avar, N., Mohammadfam, I.J.H. and Journal, t.E., "Modeling propane leak in an industry publication", Vol. 20, (2012).

Persian Abstract

چکیده

در این مقاله ارزیابی ریسک مخزن بنزین شرکت ملی پخش فرآورده های نفتی منطقه ساری به کمک نرم افزار PHAST با توجه به داده های محیطی و فرآیندی واحد مورد ارزیابی قرار گرفته است. مدل سازی پیامدهای ناشی از سناریوهای مختلف مانند نشتی کوچک ، نرخ رهایش ثابت و پارگی کامل صورت گرفت و سپس محدوده مربوط به هریک از آن ها با توجه به شدت تشعشع یا موج فشار بدست آمد و فواصل ایمن هریک از آن ها مشخص گردید. با توجه به پیامد انفجار، بدترین نتایج آن مربوط به شرایط آب و هوایی F ۲/۳ به ترتیب ۴۷۰۰، ۲۴۰۰ و ۲۳۰۰ متر می باشد. همچنین، بر اساس داده های آتش فورانی و ناگهانی، شدت تشعشع که متناظر با مرگ آنی و یا تخریب کامل تجهیزات می باشد، در شرایط آب و هوایی (F ۲/۳ و D ۴/۱)، به ترتیب در فواصل ۱۸۰ و ۱۶۰ متر اتفاق افتاد و فواصل اشتعال پذیری در این دو شرایط آب و هوایی به ترتیب ۱۰۵۲۰ و ۴۵۰ متر بوده است. در ادامه، با ترکیب شدت این حوادث با توزیع جمعیت و احتمال رخداد آن ها سطح ریسک این مخازن تعیین شد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Presenting a New Method for Earthquake Relief Center Location Allocation Based on Whale Optimization Algorithm

A. Zarepor Ashkezari[a], H. Mosalman Yazdi*[b]

[a] Department of Civil Engineering, Yazd Branch, Islamic Azad University, Yazd, Iran
[b] Department of Civil Engineering, Maybod Branch, Islamic Azad University, Maybod, Iran

PAPER INFO

Paper history:
Received 14 November 2020
Received in revised form 11 December 2020
Accepted 01 January 2021

Keywords:
Disaster Management
Earthquake
Relief Center
Heuristic Algorithms
Whale Optimization Algorithm

ABSTRACT

Despite the regulations set for the reinforcement of structures, many buildings in the world are vulnerable to earthquakes. Local governments try to be prepared to cope with this possible crisis through establishing earthquake relief centers. Considering the budget provision for creating n relief centers in a certain region, the main problem is yet in which place these centers should be constructed in order to achieve the highest speed and quality in rescuing after an earthquake. The enormous number of points that need relief, and of the many locations that could be candidates for constructing a relief center have caused this problem to be considered as an NP-Complete problem. In this article, there is a focus on solving the location allocation of the earthquake relief center problem. In order to find a reasonable solution, Whale Optimization Algorithm has been used. Classic Whale functions have been modified for this research dedicatedly. Results of the algorithm implementation and its execution on the map of region 1 of the city of Tehran show that with 9 relief centers, the average distance between each point and the center is roughly equal to 760 meters, showing almost 1.9% optimization compared to the best recent articles.

doi: *10.5829/ije.2021.34.04a.03*

1. INTRODUCTION

Disasters affecting human societies increase almost every year. Correct management of these disasters could have a great role in preserving human life. Disaster management is divided into 4 main phases. These phases consist of preventing disaster, preparation for the occurrence, responding properly to the disaster and reconstructing afterwards [1]. Among these 4 phases, proper response is the only one that runs immediately and very shortly after the disaster. Proper response to disasters includes all the work that is done immediately after a disaster in order to save the life of civilians and protect their property [2]. In other words, a proper response mission is somehow simultaneous with the disaster. Therefore, time and human resources management are very important in it. Cooperation and coordination between various parts is also one of the essential requirements for succeeding in the response phase [3,4]. This phase of disaster management has been focused on in this article.

In order to succeed in responding to disaster, conducting two sets of measures is necessary. The first set of tasks includes supplying suitable hardware equipment for providing relief after disaster. This equipment includes food, fire prevention equipment, safety equipment and medical equipment. The second set of tasks includes decisions that must be taken in advance in order to provide the best relief when responding. These decisions consist of location allocation of relief centers and shelters, highlighting relief paths and careful planning for commuting [4-8].

One of the painful disasters that threatens many people's lives is the earthquake. For reducing an earthquake effect, all disaster management phases should be applied. One of the most important effects of earthquakes is mass destruction of buildings in a region.

*Corresponding Author Email: HAMosalmanYazdi@gmail.com (H. Mosalman Yazdi)

Please cite this article as: A. Zarepor Ashkezaria, H. Mosalman Yazdib, Presenting a New Method for Earthquake Relief Center Location Allocation Based on Whale Optimization Algorithm, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 769-781

A major challenge in responding to earthquakes is location allocation of temporary relief centers in the earthquake zone. Relief centers must be placed in locations in which all points can be serviced with maximum speed. This problem is known as the Location Allocation or LA problem. The main objectives in the LA problem are reducing "relief providing time" and "rescuer dispatching cost" as much as possible [9].

The LA problem is a hard and complicated problem that doesn't have any solution in polynomial time. Having enormous points and conditions has caused finding the best solution to be impossible through classical and structural algorithms. Various complexities have been defined for this problem in different articles. In some articles, the problem of objective incompatibility has been mentioned. In other articles, having too many points has been talked about. A number of articles consider objective function complexities and limitations as the main reason of difficulty in this problem. Moreover, uncertainty and high volume of data are a couple of reasons for further complicating the problem [10,11].

As mentioned above, a problem in such a level of hardness and complexity cannot be solved by means of a classical algorithm, hence this problem, like any other NP-Complete problem, is solved by optimization methods. In optimization methods that are implemented based on heuristic algorithms, there is not only one definite answer. Moreover, the produced solution is not necessarily the optimum answer. However, these kinds of algorithms produce a certain solution in each execution that is close to the optimum answer in an acceptable extent. Therefore, the quick and reliable output of an optimization method can be an appropriate basis for the system's decision makers [12].

Nowadays, using metaheuristic methods in hard problem optimization has become common more than ever. The reasons of using these algorithms can be mentioned in a few main points: The first reason is that these algorithms are based on simple concepts and their implementation is rather easy. The second reason is that they don’t need precise information about classical solutions of the problem. The third reason is that they can easily cope with local optimums. Metaheuristic algorithms are inspired by natural and physical phenomena. Take genetic-based algorithms for example, or algorithms that have been designed by inspiration from gravity. Another group of algorithms are particle based methods [13]. These algorithms have been inspired by mass movement of particles for hunting or reaching a certain goal. The most famous particle-based algorithm is PSO. This algorithm has been inspired by the hunting process in birds. In solving problems by this method, each particle is a random solution for the problem. By means of a certain function called the “fitness function,” closeness of each solution to the best answer is calculated [14]. One of the newest particle-based algorithms is the Whale Optimization Algorithm or WOA [13].

In this article, WOA was used for solving the LA problem. Since location allication is a NP-hard problem and WOA has never been used in it, one of the motivations of this research is modifying the WOA algorithm in order to solve the problem. Thanks to characteristics and novelty of this algorithm it was anticipated that the results would be satisfactory. According to this, WOA was applied to GIS information of a group of points in a certain map in order to find the best locations for constructing relief centers. The proposed method was evaluated by means of region 1 of the city of Tehran. The main goal of this article is adjusting WOA in order to find the best locations for earthquake relief centers. The proposed method was assessed and then compared with previous studies.

In this project, all of the points were extracted from a GIS map. In this method, all points in a map could be a candidate to become a relief center. Due to some managerial reasons, if just some of these points are candidates to become relief centers, there won't be any trouble to the whole method, but the problem space will be more limited and its solving will become even easier. After determining the points, each of which could be a relief center, WOA is applied to them. This algorithm was dedicatedly modified and adjusted to solve the LA problem. The algorithm parameters were calibrated by a simple hypothetical map. After calibration, the proposed method was evaluated by a complicated map and a real map from region 1 of Tehran.

The rest of the article is organized as follows. The second section reviews previous work done in this area. The third section presents an overall introduction of WOA. In the fourth section, the proposed algorithm was introduced thoroughly and all of its parts have been explained. In the fifth section, the proposed method was implemented and evaluated from many aspects. In the last section, there is a summary, conclusion and some suggestions for future studies.

2. PREVIOUS STUDIES

For the first time, the LA problem was formulated in 1964. In this formulation, complexities of the problem, like getting stuck in local optimum and lack of convergence in objective function, were demonstrated. In the same research, efficiency of heuristic functions in solving the LA problem was highlighted [15,16]. After defining the problem, much research was done to solve it. A method was proposed by Badri [17] to solve the LA problem in a general case. In this method, a combination of the Analytic Hierarchy Process and Goal Programming was used.

Wesolowsky et al. [18] solved a special version of the LA problem. In this version, there is an ability to move facilities in certain periods of times. For solving the problem, two different methods are presented. The first one was implemented by Mixed-Integer Programming and is suitable for small-scale problems. In the second, Dynamic Programming was used. By means of the second method, large-scale problems can be solved. In various articles, some other methods, like Fuzzy Logic and the Stochastic Model were used to solve the problem as well [19,20].

After introducing and studying some of the solutions for LA, those versions of LA will be reviewed that are used for location allocation of earthquake relief centers. Due to the considerable importance of the problem, much research was conducted. Some algorithms have used direct optimization and some others have solved the problem by means of metaheuristic methods. First, direct optimization methods will be introduced. Berladi et al. [21] have developed a probabilistic algorithm for finding optimum location of relief centers. In this algorithm, relief centers are located in a non-deterministic environment. Another method was proposed in 2007 that uses a multi-dimension model for locating relief centers. In this method, parameters like cost, response time and responsibility have been considered [22]. Hooshangi and Alesheikh [23] defined a new algorithm for a disaster zone that assigns tasks to relief centers.

Another approach used for the LA relief center problem concerns metaheuristic methods. Shortages and limitations in classical methods lead researchers to use metaheuristic algorithms. These algorithms have a high ability in searching for the problem space and solving hard problems. A research conducted by Yi and Kumar [24] used PSO to solve the problem. In this research, a new method was proposed to organize the rescue mission. The mission was divided to two phases. In Phase 1, relief paths are characterized, and in the second phase a material distribution pattern was determined [24]. Ghasemi and Khalili-Damghani [25] proposed a robust method for pre-disaster location allocation inventory planning. Their method is based on a mathemathical method which can be optimized by meta-heuristic methods [25]. In another paper, Ghasemi et al. [26] proposed a method multi-period multi-vehicle location allocation model for earthquake evacuation planning. They combined an exact model with a meta-heuristic model to tackle the problem. They applied their method to the Tehran city (Iran) [26]. Paul et al. [27] presented a robust method for location allocation network design. They used an optimization model to reach a better solution. Their method was checked by the data of the Northridge in the California, USA [27]. Saedian et al. [2] focused on a new method for location allocation of relief centers. In their research, GA and BA algorithms have been used. In this method, parcels in a GIS map have been used as input of the algorithm. The algorithm has been evaluated on region 1 of Tehran [2]. Based on parcel data and region 1 of Tehran, another algorithm has been developed as well. It is a hybrid method that has combined GA and BA with clustering methods and TOPSIS [28].

Metaheuristic methods are used to solve a wide range of problems. The main ability of these methods is searching problem space and finding reasonable answers. Therefore, in problems which do not have a classical solution, the best approach is using metaheuristic methods [29]. There are various kinds of metaheuristic methods. Some of them, like GA, have been designed based on genetic science [30]. Some other algorithms, like SA, have been invented by inspiration from physical phenomena [31]. A well-known group of metaheuristic methods are particle-based algorithms. There is a wide range of particle-based algorithms. Algorithms like PSO, BA, GOA and ACO are all particle-based [14,32,33,34]. One of the newest algorithms that was proposed in recent years is WOA or Whale Optimization Algorithm. WOA gets inspiration from mass hunting of humpback whales [13]. In this article, WOA has been used for the earthquake relief center location allocation problem.

WOA was used to solve many problems in recent years. Mirjalili et al. [34] have used WOA for feature selection in data mining. They combined WOA with SA to manage to select the best features [35]. In other research, WOA was used to optimize the weight of edges in ANN [36]. Moreover, WOA has been applicable in image processing. For example, in 2017 a new segmentation method was designed based on WOA [37]. WOA is also usable in multi-objective problems. Wang et al. [38] proposed a new method for wind speed forecast based on WOA. Researchers have also used WOA in scheduling problems. In 2018, an algorithm based on WOA was proposed to solve the Flow Shop problem [39]. In this article, WOA was applied to solve the earthquake relief center location allocation problem.

Selecting WOA to solve the LA problem has some sensible reasons. The first reason is that WOA is a rather new algorithm that has never been used for the LA problem so far. Therefore, using WOA can lead to new results for the problem. The second reason is that in recent years, WOA was used in many areas of science and has had good results as well. These successes have motivated authors of the article to use WOA in the LA problem. The third reason goes back to inherent characteristics of WOA. WOA has a good convergence speed and can truly transmit from the exploration phase to exploitation phase. It has a high ability to escape from local optimum. Given the above reasons, it feels that WOA can be an effective tool to solve the LA problem.

3. WHALE OPTIMIZATION ALGORITHM

Whale Optimization Algorithm was created based on the mass movement of humpback whales during hunting. It is classified as a particle-based algorithm. The performance of a whale in WOA is just like a particle in PSO. Two different types of movement toward prey have been considered for whales. One of them is Spiral, in which a whale goes toward prey through a spiral path. Another movement is called Shrinking. In Shrinking, every whale goes directly toward prey. In fact, group Shrinking causes the mass of whales to contract. Each whale is a random solution for the LA problem. Apart from Spiral and Shrinking, another function, called Searchprey also exists. This function adds some of the genetic-based algorithm's characteristics to WOA. Basic WOA is a continuous algorithm. To use this algorithm in the LA problem, which is a discrete problem, Spiral, Shrinking and Searchprey functions have been redefined with a new performance [13].

The values of a, A, l and p, are parameters of this algorithm. The parameter p has a value in [0,1] and through it, either Spiral or Shrinking will be selected to run. The parameter a varies from 2 to 0 in order for parameter A to be made. The parameter A transmits the algorithm from the exploration to exploitation phase. The parameter l is for rotation angle that has been explained in the next section. Figure 1 illustrates the execution process of Whale Optimization Algorithm.

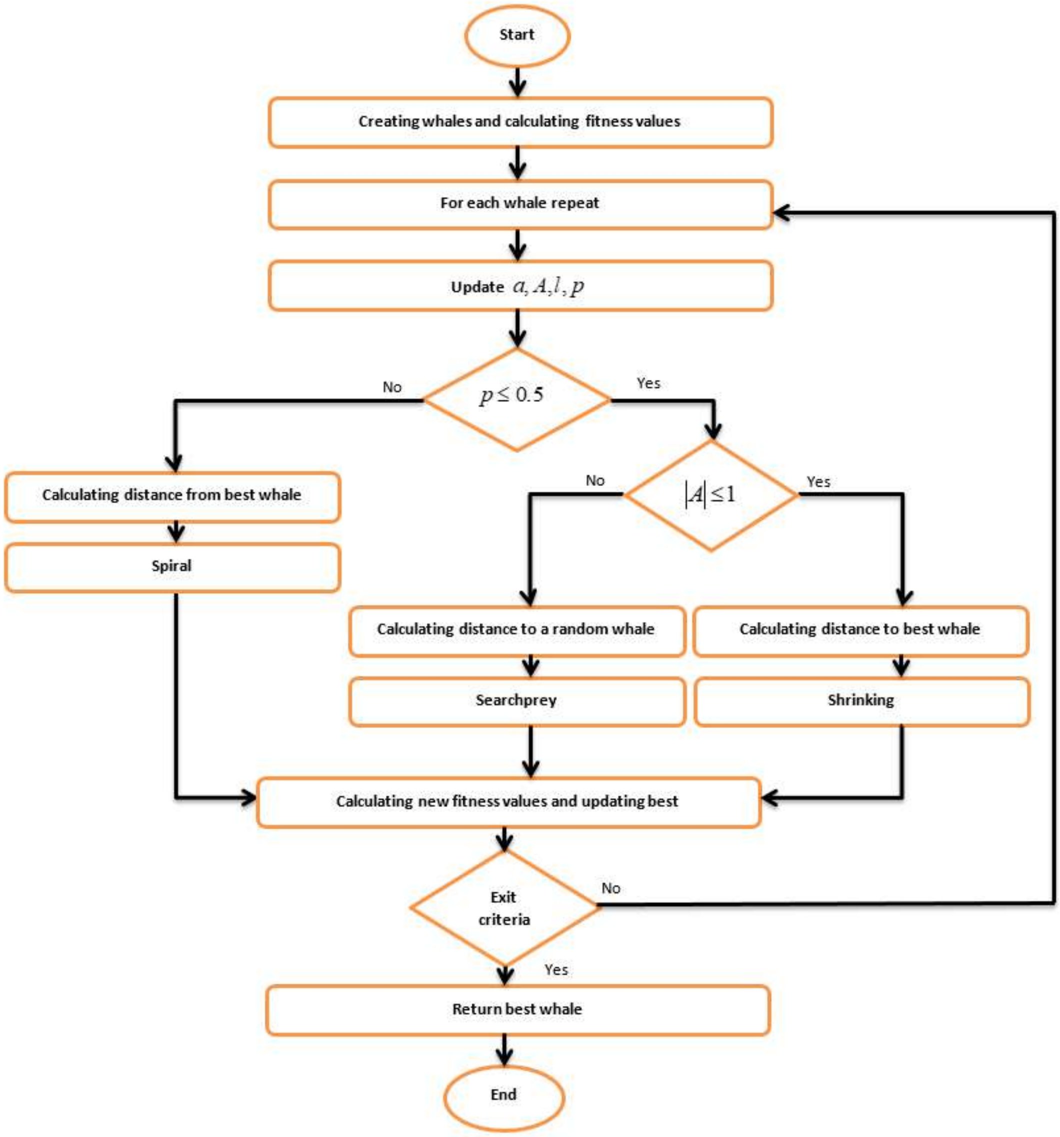


Figure 1. Structure of WOA in the proposed method

4. THE PROPOSED METHOD

In this section, the proposed method is explained thoroughly. Before explanation of the proposed method, an overall schematic flow diagram is illustrated in Figure 2. As it is shown in Figure 2, GIS maps of a region enter the algorithm as an input. Another implicit input that is considered number of relief centers. Following, more details of the designed algorithm are explained.

4. 1. Structure of a Whale In the proposed method, each whale represents a solution for locating relief centers. An analogy can be drawn between a whale in WOA, a particle in PSO and a chromosome in GA. Initially, whales are created randomly, and in the execution process, they are optimized by designed functions. Before describing designed WOA, whale structure and its design pattern should be explained. Data structure of each whale is a dynamic array whose length is specified by the number of relief centers. Each index of this array consists of a certain coordinate that indicates the proposed location for a relief center. Table 1 shows an example of whales that was created for a problem with 5 relief centers. In this whale, the first index indicates location of the first center; the second index indicates location of the second center; and so on.

4. 2. Objective Function The objective function or fitness function is a function that indicates worthiness

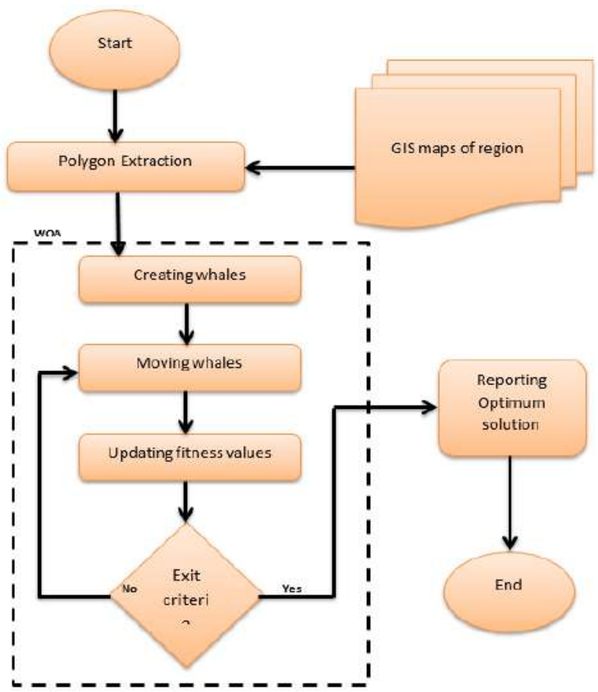


Figure 2. Overall schema of the proposed method

TABLE 1. An example of whales with 5 centers

Whale	$Center_1$	$Center_2$	$Center_3$	$Center_4$	$Center_5$
X	1287	755	1065	356	790
Y	280	799	1899	521	1572

of each whale. By calculating each whale's fitness value, it could be possible to move whales and optimize the solutions. Before explaining fitness function, it is worth mentioning that parcels are assigned to relief centers by the Euclidean method. In other words, the duty of servicing to a parcel should be done by the nearest relief center. Logically, a shorter distance between parcels and centers leads to faster relief. Therefore, the first notion to design a fitness function comes from the Euclidean distance between parcels and centers. This function is calculated by Equation 1. In this equation, $whale$ is a whale whose fitness value is calculated. The variable n indicates number of parcels on the map. $parcel_i$ is a parcel on the map. $center_i$ shows the assigned center to $parcel_i$. This function has been designed in a way that lower value of fitness indicates higher quality of whale. In fact, whale optimization problem searches for lessening fitness value.

$$fitness(whale) = \frac{\sum_{i=1}^{n} distance(parcel_i, center_i)}{n} \tag{1}$$

There is a major weakness in the fitness function of Equation (1). The major weakness is that load-balancing in centers has been overlooked. The only factor that affects this function is the distance between parcels and points. This property causes too many parcels to be assigned to a single center in crowded areas. On the other hand, in dispersed areas, the number of parcels in each center would be very low. This weakness leads to overhead in some centers and waste of facilities in some other. To solve the problem, a penalty function has been defined. This function tries to assign a grade to each whale based on load-balancing. The more balance in centers we have, the better grades would be produced. The best state is a completely balanced distribution of parcels. Equation (2) shows the penalty function. In this equation, variable n indicates the number of parcels on the map. $parcel_i$ is a parcel on the map. $center_i$ shows the assigned center to $parcel_i$. And k is the number of centers.

$$penalty(center_i) = \begin{cases} 1 & parcels_i < \frac{n}{k} \\ 0 & O.W. \end{cases} \tag{2}$$

Now, the fitness function of Equation 1 can be improved by the penalty function. Equation 3 shows the new fitness function. In this function, both factors of closeness to centers and load balancing will be effective.

$$fitness(whale) = \frac{\sum_{i=1}^{n} distance(parcel_i, center_i)}{\sum_{j=1}^{k} penalty(center_j)} \quad (3)$$

4. 3. Definition of Distance One of the most important concepts in WOA is the distance between whales. In basic WOA, distance is calculated by a subtraction. For this problem, the distance should be redefined. In the proposed algorithm, the distance between two whales is in the form of an array. The length of the array equals length of whale. Each index contains a number that is calculated by Euclidean distance. The distance of a center from the closest center in the other whale is considered as a value in the array. After calculating this value for all array cells, we will have the complete distance array. Algorithm 1 shows the computation of the distance array.

Algorithm 1 : whaleDistance

$Input : w_a , w_b$
$Output : disVector, nearest$

$k = length(w_a)$
$disVector = array\ with\ length\ k, initialize\ by\ +\infty$
$selected = array\ with\ length\ k, initialize\ by\ zero$
$nearest = array\ with\ length\ k, initialize\ by\ zero$
$for\ i = 1\ to\ k$
  $for\ j = 1\ to\ k$
    $if\ distance(w_a(i), w_b(j)) < disVector(i)\ and\ not(selected(j))$
      $disVector(i) = distance(w_a(i), w_b(j))$
      $tmp = j;$
      $nearest(i) = j;$
  end
  $selected(tmp) = 1;$
end

Figure 3 shows an example of two whales. The red points show the centers of the first whale, and the black points show those of the second. Apart from the sequence of centers, the distance of each center in the first whale is calculated with the closest center in the second.

4. 4. Whale Move Function Whales move through three different functions. These functions have been rewritten for earthquake relief center LA problem. The functions have been designed in a way that carries the concepts of basic WOA. This section describes these functions.

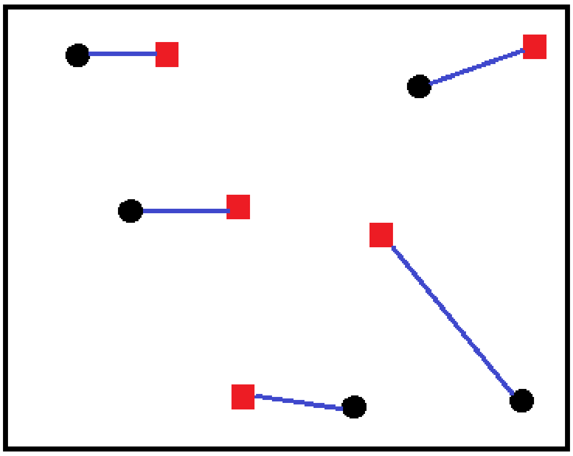


Figure 3. Distance of red and black whales by blue line

4. 4. 1. Spiral In some states of the algorithm, the whales should spiral around the best whale. The Spiral function simulates this movement. Spiral is a controlled and smooth movement toward the prey. For Spiral movement, first, the distance array should be calculated. Then a random variable in [-1,+1] , called l, is produced. By means of $2\pi l$, the rotation angle is calculated. Assuming that centers of the whale w_1 should move toward the best whale, for each pair of centers, a random point is produced on an imaginary line of w_1 to the best whale. Then, the point turns clockwise with a random point as center and $2\pi l$ as rotation angle. In this way, Spiral is a function with a kind of rotation toward the prey. Figure 4 illustrates an example of Spiral. The rotation angle is assumed to be 45 degrees. The red point is the random point between two whales, and the green point is destination of rotation. This operation should be done for all centers. To simplify the figure, just one move has been illustrated.

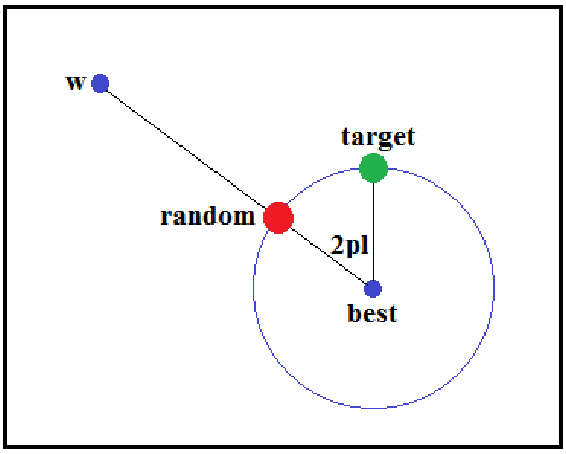


Figure 4. An example of Spiral

Algorithm 2 shows the pseudocode of the Spiral function. In this pseudocode, $point$ is the random point, $destination$ is the new location of a center and w is the new whale after Spiral movement.

Algorithm2: Spiral

Input : $\boldsymbol{w_1}$: *a whale for spiral* , **best** : *the best whale*
Output : $\boldsymbol{w}$: *spiraled whale*

$l = a\ random\ number\ in\ [-1,+1];$
$\theta = 2\pi l;$ // *rotation angle*
for $i = 1$ *to* $len(w)$
  $point = a\ random location\ on\ line(w_1, best);$
  $destination = \theta$ *sized roundClock rotation centered by best*;
  $w(i) = destination;$
end

4. 4. 2. Shrinking The shrinking function simulates directed movement toward the best whale. This function has two parameters. The first one is a typical whale and the second one is the best whale. As a result of Shrinking, a typical whale moves toward the best whale. For the whale that is supposed to move, first the distance array should be calculated. Then random points are produced on the imaginary line between centers of the whale and the best whale. At the end, the centers are transferred to random points. The big difference between Shrinking and Spiral is that in Shrinking, we don't have rotation angle. In fact, Shrinking simulates a direct and fast move. Algorithm 3 shows the pseudocode of Shrinking.

Algorithm3: Shrinking

Input : $\boldsymbol{w_1}$: *a whale for spiral* , **best** : *the best whale*
Output : $\boldsymbol{w}$: *shrinked whale*

for $i = 1$ *to* $len(w)$
  $destination = a\ random location\ on\ line(w_1, best);$
  $w(i) = destination;$
end

4. 4. 3. Searchprey The Searchprey function adds the ability of finding new solutions to the proposed method. This function does not need the best whale and works with typical whales. Execution of this function leads to better exploration of the problem space. Algorithm 4 shows the pseudocode of Searchprey.

As it can be seen in Algorithm 4, three different functions are used in the algorithm. One of the functions is Shrinking, and the others are the Join and Randomwalk function. These two functions are similar to mutation and crossover in genetics.

4. 5. Parameter a Parameter a is one of the most important values in WOA. The formula of

Algorithm4 : searchprey

Input : $\boldsymbol{w_1}$: *a whale for shrinking* , $\boldsymbol{w_2}$: *a randomly selected whale*
Output : $\boldsymbol{w}$: *a new whale*

$distance = whaleDistance(w_1, w_2);$
$tmp = a\ random\ number\ in\ [0,1];$
if $tmp < 0.25$
  $w = shrinking(w_1, w_2);$
elseif $tmp < 0.25$
  $w = join(w_1, w_2);$
else
  $w = randomwalk(w_1);$
end

producing a creates descending values in [2,0]. In other words, in the beginning of the algorithm, a equals 2, but its value gradually decreases. Eventually, a would be a number around zero. Equation 4 shows the production formula of a. In this equation, k is maximum number of whale movements, and i is the current number of whale movements.

$$a = 2 \times \left(\frac{k-i+1}{k}\right)^2 \tag{4}$$

Assuming $= 1000$, Figure 5 shows the changes in a with blue color. Given this changing pattern in, the likelihood of Searchprey at the beginning of the algorithm and that of Shrinking at the end, are very high. In fact, the proposed algorithm produces new random solutions in the beginning and optimizes the produced solutions at the end.

In order to create a little bit of a random condition in the exploitation phase, a new parameter, called A is created by a. Most of the times, the value of A is close to a. The last decision about Shrinking or Searchprey will be made by A. Equation 5 shows the calculation of A. r is a random number with normal distribution in [0,2]. Figure 5 shows the changes in A with red color.

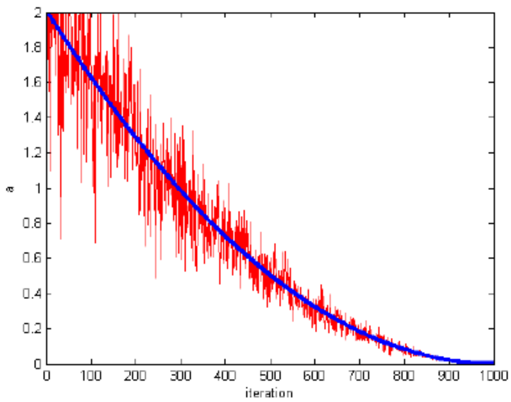


Figure 5. Changing procedure in a for $k = 1000$

$$A = 2ar - a \tag{5}$$

5. EVALUATION OF THE PROPOSED METHOD

In this section, the proposed method has been implemented and evaluated. Matlab software has been used for implementation. After calibration, the proposed method was tested by hypothetical maps and region 1 of Tehran's map. Input of the algorithm is parcels of polygon that can be extracted from GIS maps.

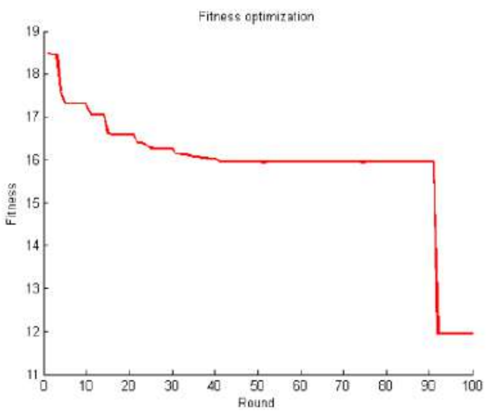
5. 1. Calibration

The proposed method has multiple parameters that could have various values. Parameters like number of movements and number of whales can considerably affect the algorithm. In this section, by means of an even map whose solution is obvious, the optimum value of parameters has been found. The scale of the map is 250 in 250. Parcels are distributed in chess form at a distance of five. In fact, the map has 2500 parcels. The optimum solution for this map is segmenting the map and locating each relief center in the center of segments. In this section, this problem has been solved with four centers. The number of whales and movements increase gradually in order to find the best value of parameters. Results of various runs have been listed in Table 1. The number of whales are 20, 30 and 40. Also, the number of movements is between [60, 100]. As it can be seen in Table 1, the best answer is the 9th row that has 40 whales and 100 moves. Executions with movements below 100 and whales below 40 have produced weaker solutions. In terms of runtime, the 9th row is the worst. The reason is obvious; to create more whales and more movements, more time is needed. In fact, by spending more time, we can explore the problem space more carefully.

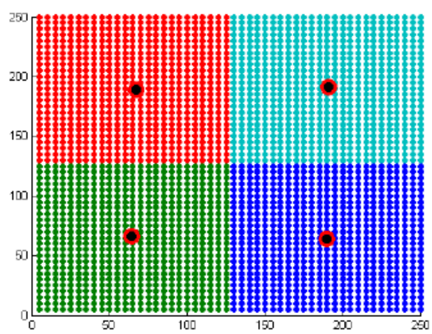
For better evaluation of the optimizing process in the 9th row of Table 2, the chart of fitness is given in Figure 6a. As it can be seen, optimization in the first rounds is satisfactory. At around the 45th round, the optimization was stopped and the best solution did not change until the 90th round. In final rounds, by a slight change, considerable optimization has been achieved. Figure 6b shows the location of centers in the final solution. This figure highlights that the optimum answer has been found.

TABLE 2. Results of the proposed method with different parameters

Row#	Whale#	Move#	Center1	Center2	Center3	Center4	fitness	Time(m)
1	20	60	512	580	669	739	21	2.5
2	20	80	643	599	572	686	19.5	4
3	20	100	602	535	711	652	19	5
4	30	60	699	605	614	582	19	6
5	30	80	615	641	610	634	18	7
6	30	100	637	620	618	625	16.5	8
7	40	60	621	625	630	624	16	5.5
8	40	80	621	625	630	624	16	7
9	40	100	625	625	625	625	12	9.5



(a) Fitness chart



(b) Location of centers

Figure 6. results of calibration (checkered environment with 4 centers)

5. 2. Evaluation with Hypothetical Maps In the hypothetical map, there are 2000 parcels in a 2000 in 2000 coordinate. Concentration of parcels in location (0,0) and around it is more, and by getting away from this point, the density of parcels would be less. It is expected that relief centers are more likely to be in congested areas. Figure 7 shows this map.

Location allocation in this map has been solved with 40 whales and 100 movements. The number of centers equals six. Evaluation of results shows that centers in congested areas are denser. Because of the ability of the fitness function, despite the fact that the map is heterogeneous, location allocation is rather balanced. Table 3 shows the results of the algorithm for both fitness functions. The first row belongs to Equation (1),

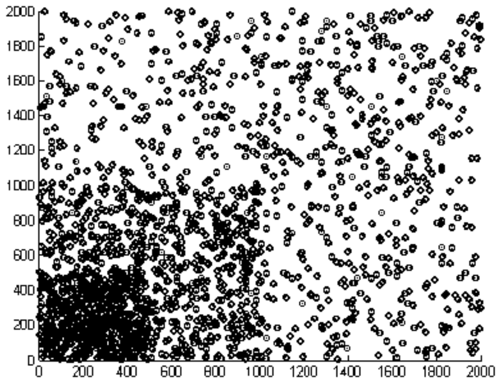


Figure 7. Heterogeneous map

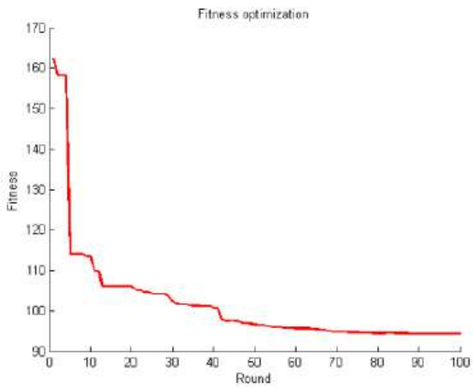
and the second to Equation (3). Comparison of results shows that with a slight increase in distance, we managed to achieve better load balancing. In the first row, the biggest number is 728, and the smallest number is 176. Whereas, in the second row, the biggest number is 590 and the smallest number is 176. Decreasing distance between the two numbers shows that load balancing has been considered. Low average in the first row shows that Equation (1) focuses on Euclidean distance.

Figure 8a illustrates the fitness chart of this problem. The optimization process indicates that the proposed algorithm has passed a sensible path to reach the final solution. At the beginning, the optimization speed is very high, whereas at the end, optimization is rare and in low speed. This kind of progress is completely matched with evolutionary algorithm's philosophy. Figure 8b shows the location of centers in the final solution. It can be easily seen that the distance between the centers in disperse parts of the map is more.

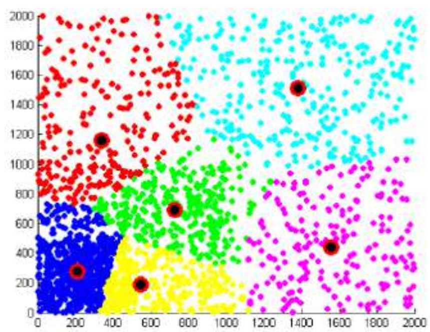
5. 3. Evaluation with Tehran's Map Iran is a country which has suffered from earthquakes quite a lot of times. Meanwhile, Tehran, which is the biggest and most populated city in the country, is an earthquake-prone city [40]. Region 1 of Tehran, as one of the most earthquake-prone regions of the city, has a population of nearly 400000, and covers an area of more than 64 square km. To locate relief centers, parcel's location should be clear. Therefore, among the layers of the GIS map, the parcel layer has been used. Figure 9 illustrates the parcels. This map has roughly 35000 parcel [2].

TABLE 3. Results in the heterogeneous map

fitness	Center1	Center2	Center3	Center4	Center5	Center6	Avg_dis
(1)	201	288	728	354	176	253	272
(3)	251	318	317	317	590	207	283.5



(a) Fitness chart



(b) Location of centers

Figure 8. Results of heterogeneous map with 6 centers

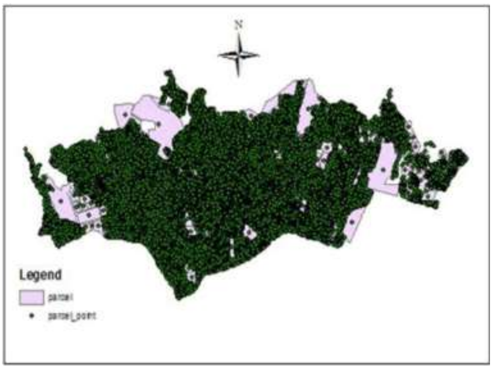


Figure 9. Parcels in region 1 of Tehran

In this section, the proposed algorithm has been applied to region 1 of the Tehran map. The number of whales is 40, and the number of movements is 100. Region 1 of Tehran has roughly 34000 parcels. 9 relief centers are supposed to be created, then parcels distributed among them. Figure 10a shows the fitness chart for 100 movements. Figure 10b illustrates the final location of relief centers. This result belongs to the fitness function that considers load balancing. Parcels of each center are distinguished by a certain color, and the relevant center can be seen in the center of parcels.

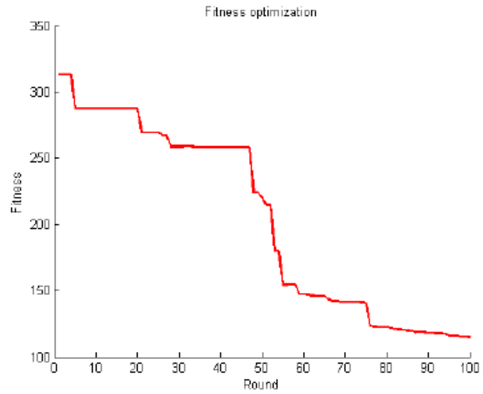
For better evaluation, details of the results have been listed in Table 4. The first row belongs to load balancing fitness function, and the second row shows the results of distance fitness function. The last column includes the average distance of parcels with their center. Comparison of values in the last column shows that the distance fitness function is better. However, focusing on the number of parcels in each center shows that the load balancing method has a more reasonable distribution of parcels. In this case, no center is overloaded, whereas, in distance function, there is a considerable distance between number of parcels. Center 1 has 4732 parcels, while center 8 has 2746.

5. 4. Stability of the Algorithm

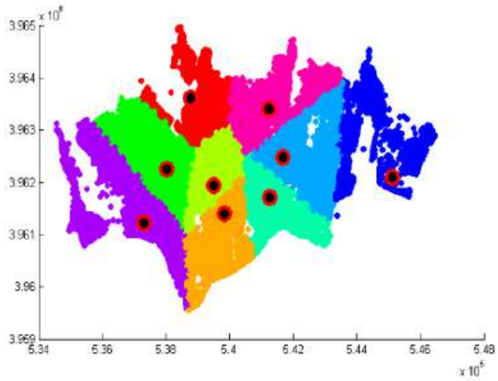
The proposed method has been designed based on an evolutionary algorithm. The main difference between this algorithms and classical methods is that optimization in evolutionary algorithms is quite random. In other words, in evolutionary algorithms, after each run, we have different solutions. Stability is one of the challenges in this algorithm. To assess stability, the proposed method was run for 35 times. Both fitness functions have been used. Figure 11 illustrates the results. The average of results is almost close to the best solution. For example, in Tehran's map with load balancing function (Figure 10c), the average is 933 and standard deviation is 13, whereas the best answer is 918. In distance function (Figure 10d) the average is 782 and standard deviation is 17, whereas the best answer is 760.

5. 5. Comparison

One of the maps that the proposed method has been evaluated with is the Tehran map's region 1. Since Saeidian et al. [2] has used region 1 of the Tehran map as well, we can compare the proposed algorithm with literature [2]. Table 5 shows the comparison results. The first column includes average distance of parcels with their centers. The second column lists the total distance of providing



(a) Fitness chart

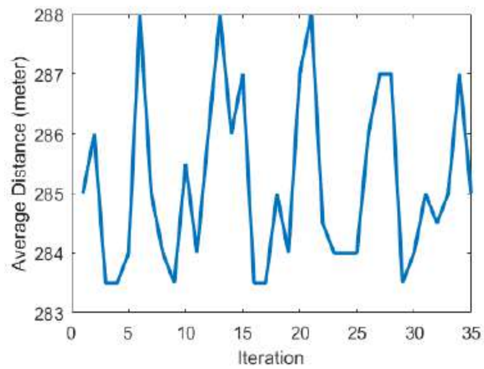


(b) Location of centers

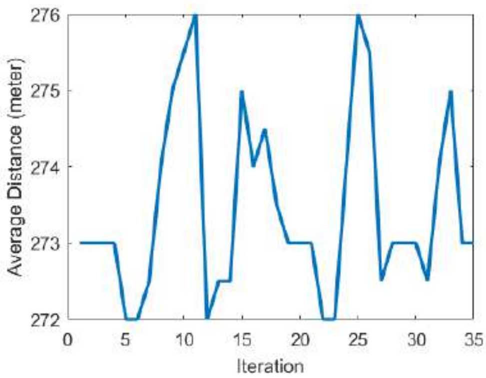
Figure 10. Result of the Tehran map (region 1) with 9 centers

TABLE 4. Details of results for Tehran (region 1)

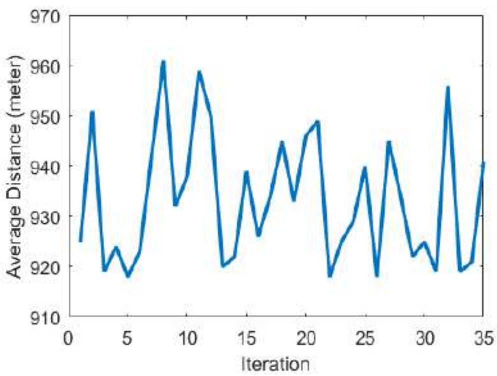
fitness	Center1	Center2	Center3	Center4	Center5	Center6	Center7	Center8	Center9	Avg_dis (meter)
Balance(3)	3757	3855	350	3922	3654	3836	3893	3959	3662	918
Mean(1)	4732	3918	4956	2873	3157	3491	4284	2746	3941	760



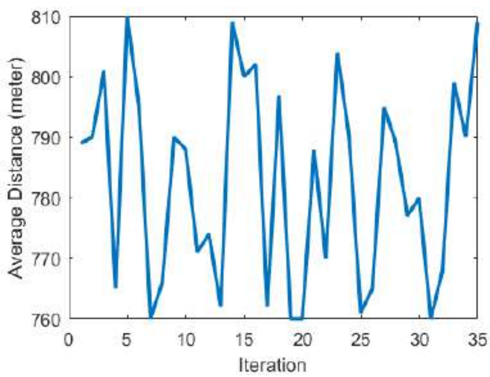
(a) Load balancing function with the hypothetical map



(b) Distance function with the hypothetical map



(c) Load balancing function with the Tehran map



(d) Distance function with the Tehran map

Figure 10. results of 35 run

TABLE 5. Comparison

fitness	Avg_dis (meter)	KM
Balance (3)	918	31.3
Mean (1)	760	25.9
[2]	775	26.4

relief. As it can be seen, the distance function (row 2) produced a better solution than literature [2] (760 vs. 775). However, by the load balancing function (row 3) the answer is weaker (918 vs. 775). Increasing in average distance is because the function wants to provide load balancing.

6. CONCLUSION

In this article, a new method has been proposed to locate relief centers. The proposed method is based on WOA and can solve the problem on any map. This method has been evaluated with many maps and eventually applied to Tehran's map. A new fitness function has been designed that has the ability of load balancing. The average distance of relief centers with parcels is 760 meter which is 1.9% better than previous work. In load balancing function the average distance has increased to 918, however the method did this sacrifice in order to improve the load balancing. Advantages of the proposed method are the following:

- Quick convergence and local optimum avoidance
- Using a new objective function for load balancing
- Dynamic of the proposed method in location of centres
- Ability of combining the proposed method with other algorithms

Based on limitations of the proposed method, the future works may include the following:

- Using other layers of maps, such as roads
- Considering the level of relief in calculations
- Considering the building vulnerability parameter in calculations

7. REFERENCES

1. Beiki, H., Seyedhosseini, S. M., Ghezavati, V. R., and Seyedaliakbar, S. M., "A location-routing model for assessment of the injured people and relief distribution under uncertainty", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 7, (2020), 1274-1284. https://doi.org/10.5829/IJE.2020.33.07A.14
2. Saeidian, B., Mesgari, M. S., and Ghodousi, M., "Evaluation and comparison of Genetic Algorithm and Bees Algorithm for location–allocation of earthquake relief centers", ***International***

Journal of Disaster Risk Reduction, Vol. 15, (2016), 94-107. https://doi.org/10.1016/j.ijdrr.2016.01.002

3. Kondaveti, R., and Ganz, A., "Decision support system for resource allocation in disaster management", In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (2009), 3425-3428, https://doi.org/10.1109/IEMBS.2009.5332498
4. Rolland, E., Patterson, R. A., Ward, K., and Dodin, B., "Decision support for disaster management", ***Operations Management Research***, Vol. 3, No. 1-2, (2010), 68-79. https://doi.org/10.1007/s12063-010-0028-0
5. Altay, N., and Green III, W. G., "OR/MS research in disaster operations management", ***European Journal of Operational Research***, Vol. 175, No. 1, (2006), 475-493. https://doi.org/10.1016/j.ejor.2005.05.016
6. Lin, Y. H., Batta, R., Rogerson, P. A., Blatt, A., and Flanigan, M., "Location of temporary depots to facilitate relief operations after an earthquake", ***Socio-Economic Planning Sciences***, Vol. 46, No. 2, (2012), 112-123. https://doi.org/10.1016/j.seps.2012.01.001
7. Najafi, M., Eshghi, K., and Dullaert, W., "A multi-objective robust optimization model for logistics planning in the earthquake response phase", ***Transportation Research Part E: Logistics and Transportation Review***, Vol. 49, No. 1, (2013), 217-249. https://doi.org/10.1016/j.tre.2012.09.001
8. Tzeng, G. H., Cheng, H. J., and Huang, T. D., "Multi-objective optimal planning for designing relief delivery systems", ***Transportation Research Part E: Logistics and Transportation Review***, Vol. 43, No. 6, (2007), 673-686. https://doi.org/10.1016/j.tre.2006.10.012
9. Jing, Wang, Zhu Jianming, Huang Jun, and Zhang Min., "Multi-level emergency resources location and allocation.", In 2010 IEEE ***International Conference on Emergency Management and Management Sciences***, 202-205. IEEE, (2010), https://doi.org/10.1109/ICEMMS.2010.5563466
10. Aerts, J. C., and Heuvelink, G. B., "Using simulated annealing for resource allocation", ***International Journal of Geographical Information Science***, Vol. 16, No. 6, (2002), 571-587. https://doi.org/10.1080/13658810210138751
11. Zheng, Y. J., and Ling, H. F., "Emergency transportation planning in disaster relief supply chain management: a cooperative fuzzy optimization approach", ***Soft Computing***, Vol. 17, No. 7, (2013), 1301-1314. https://doi.org/10.1016/j.ejor.2005.05.016
12. Aerts, J. C., Goodchild, M. F., and Heuvelink, G. B., "Accounting for spatial uncertainty in optimization with spatial decision support systems", ***Transactions in GIS***, Vol. 7, No. 2, (2003), 211-230. https://doi.org/10.1111/1467-9671.00141
13. Mirjalili, S., and Lewis, A., "The whale optimization algorithm", ***Advances in Engineering Software***, Vol. 95, (2016), 51-67. https://doi.org/10.1016/j.advengsoft.2016.01.008
14. Kennedy, J., "Particle swarm optimization", ***Encyclopedia of Machine Learning***, (2010), 760-766. https://doi.org/10.1109/ICNN.1995.488968
15. Cooper, L., "Heuristic methods for location-allocation problems", ***SIAM Review***, Vol. 6, No. 1, (1964), 37-53. https://doi.org/10.1137/1006005
16. Cooper, L., "Location-allocation problems", ***Operations Research***, Vol. 11, No. 3, (1963), 331-343. https://doi.org/10.1287/opre.11.3.331
17. Badri, M. A., "Combining the analytic hierarchy process and goal programming for global facility location-allocation problem", ***International Journal of Production Economics***, Vol. 62, No. 3, (1999), 237-248. https://doi.org/10.1016/S0925-5273(98)00249-7
18. Wesolowsky, G. O., and Truscott, W. G., "The multiperiod location-allocation problem with relocation of facilities", ***Management Science***, Vol. 22, No. 1, (1975), 57-65. https://doi.org/10.1287/mnsc.22.1.57
19. Zhou, J., and Liu, B., "New stochastic models for capacitated location-allocation problem", ***Computers & Industrial Engineering***, Vol. 45, No. 1, (2003), 111-125. https://doi.org/10.1016/S0360-8352(03)00021-4
20. Zhou, J., and Liu, B., "Modeling capacitated location–allocation problem with fuzzy demands", ***Computers & Industrial Engineering***, Vol. 53, No. 3, (2007), 454-468. https://doi.org/10.1016/j.cie.2006.06.019
21. Beraldi, P., and Bruni, M. E., "A probabilistic model applied to emergency service vehicle location", ***European Journal of Operational Research***, Vol. 196, No. 1, (2009), 323-331. https://doi.org/10.1016/j.ejor.2008.02.027
22. Tzeng, G. H., Cheng, H. J., and Huang, T. D., "Multi-objective optimal planning for designing relief delivery systems", Transportation Research Part E: ***Logistics and Transportation Review***, Vol. 43, No. 6, (2007), 673-686. https://doi.org/10.1016/j.tre.2006.10.012
23. Hooshangi, N., and Alesheikh, A. A., "Agent-based task allocation under uncertainties in disaster environments: An approach to interval uncertainty", ***International Journal of Disaster Risk Reduction***, Vol. 24, (2017), 160-171. https://doi.org/10.1016/j.ijdrr.2017.06.010
24. Yi, W., and Kumar, A., "Ant colony optimization for disaster relief operations", Transportation Research Part E: ***Logistics and Transportation Review***, Vol. 43, No. 6, (2007), 660-672. https://doi.org/10.1016/j.tre.2006.05.004
25. Ghasemi, P., and Khalili-Damghani, K., "A robust simulation-optimization approach for pre-disaster multi-period location–allocation–inventory planning", ***Mathematics and Computers in Simulation***, Vol. 179, (2020), 69-95. https://doi.org/10.1016/j.matcom.2020.07.022
26. Ghasemi, P., Khalili-Damghani, K., Hafezalkotob, A., and Raissi, S., "Uncertain multi-objective multi-commodity multi-period multi-vehicle location-allocation model for earthquake evacuation planning", ***Applied Mathematics and Computation***, Vol. 350, (2019), 105-132. https://doi.org/10.1016/j.amc.2018.12.061
27. Paul, J. A., and Wang, X. J., "Robust location-allocation network design for earthquake preparedness", ***Transportation research part B: Methodological***, Vol. 119, (2019), 139-155. https://doi.org/10.1016/j.trb.2018.11.009
28. Saeidian, B., Mesgari, M., Pradhan, B., and Ghodousi, M., "Optimized Location-Allocation of Earthquake Relief Centers Using PSO and ACO, Complemented by GIS, Clustering, and TOPSIS", ***ISPRS International Journal of Geo-Information***, Vol. 7, No. 8, (2018), 292. https://doi.org/10.3390/ijgi7080292
29. Kahrizi, M. R., and Kabudian, S. J., "Projectiles Optimization: A Novel Metaheuristic Algorithm for Global Optimization", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 10, (2020), 1924-1938. https://doi.org/10.5829/IJE.2020.33.10A.11
30. Koza, J. R., "Genetic programming", ***Search Methodologies,*** (1997), 127-164. https://doi.org/10.1007/0-387-28356-0_5
31. Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P., "Optimization by simulated annealing", ***Science***, Vol. 220, No. 4598, (1983), 671-680. https://doi.org/10.1126/science.220.4598.671
32. Yang, X. S., "A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization", NICSO 2010, (2010), 65-74. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12538-6_6

33. Dorrani, Z., Farsi, H., and Mohamadzadeh, S., "Image Edge Detection with Fuzzy Ant Colony Optimization Algorithm", ***International Journal of Engineering, TransactionC: Aspects***, Vol. 33, No. 12, (2020), 2464-2470. https://doi.org/ 10.5829/IJE.2020.33.12C.05

34. Mirjalili, S. Z., Mirjalili, S., Saremi, S., Faris, H., and Aljarah, I., "Grasshopper optimization algorithm for multi-objective optimization problems", ***Applied Intelligence***, Vol. 48, No. 4, (2018), 805-820. https://doi.org/10.1007/s10489-017-1019-8

35. Mafarja, M. M., and Mirjalili, S., "Hybrid whale optimization algorithm with simulated annealing for feature selection", ***Neurocomputing***, Vol. 260, (2017), 302-312. https://doi.org/10.1016/j.neucom.2017.04.053

36. Aljarah, I., Faris, H., and Mirjalili, S., "Optimizing connection weights in neural networks using the whale optimization algorithm", ***Soft Computing***, Vol. 22, No. 1, (2018), 1-15. https://doi.org/10.1007/s00500-016-2442-1

37. El Aziz, M. A., Ewees, A. A., and Hassanien, A. E., "Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation", ***Expert Systems with Applications***, Vol. 83, (2017), 242-256. https://doi.org/10.1016/j.eswa.2017.04.023

38. Wang, J., Du, P., Niu, T., and Yang, W., "A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting", ***Applied Energy***, Vol. 208, (2017), 344-360. https://doi.org/10.1016/j.apenergy.2017.10.031

39. Abdel-Basset, M., Manogaran, G., El-Shahat, D., and Mirjalili, S., "A hybrid whale optimization algorithm based on local search strategy for the permutation flow shop scheduling problem", ***Future Generation Computer Systems***, Vol. 85, (2018), 129-145. https://doi.org/10.1016/j.future.2018.03.020

40. Ibrion, M., Mokhtari, M., and Nadim, F., "Earthquake disaster risk reduction in Iran: Lessons and “lessons learned” from three large earthquake disasters—Tabas 1978, Rudbar 1990, and Bam 2003", ***International Journal of Disaster Risk Science***, Vol. 6, No. 4, (2015), 415-427. https://doi.org/10.1007/s13753-015-0074-1

Persian Abstract

چکیده

با وجود قوانین تعیین شده برای مقاوم‌سازی سازه‌ها، ساختمان‌های زیادی در دنیا در برابر زلزله آسیب‌پذیر هستند. ازاین‌رو وقوع زلزله در بسیاری از نقاط جهان موجب ایجاد بحران خواهد شد. به همین دلیل یکی از سرفصل‌های مدیریت بحران، مدیریت امداد و نجات بعد از زلزله است. دولت‌های محلی با ایجاد مراکز امداد زلزله سعی می‌کنند برای مقابله با این بحران احتمالی آماده شوند. بودجهٔ محدود و هزینهٔ بالای ایجاد مراکز امداد زلزله موجب شده تا مکان‌یابی این مراکز به یک مسئلهٔ مهم تبدیل شود. با فرض تأمین بودجه برای احداث n مرکز زلزله در یک منطقه، مسئلهٔ اصلی این است که این مراکز در چه مکانی ساخته شود تا بالاترین سرعت و کیفیت امدادرسانی بعد از زلزله را داشته باشد. تعداد بیشمار نقاطی که به امداد نیاز دارند و تعداد زیاد مکان‌هایی که می‌تواند کاندیدای احداث مرکز امداد باشد موجب شده این مسئله در مجموعهٔ مسائل np-complete قرار گیرد. برای حل چنین مسائلی باید از روش‌های اکتشافی و تکاملی استفاده کرد. در این مقاله روی حل مسئله مکان‌یابی مراکز امداد زلزله تمرکز شده است. برای یافتن یک پاسخ از الگوریتم تکاملی نهنگ استفاده شده است. الگوریتم نهنگ یک الگوریتم فرااکتشافی است که با الهام از شکار دسته‌جمعی نهنگ‌های کوهان‌دار طراحی شده است. توابع کلاسیک نهنگ برای حل این مسئله به‌طور اختصاصی بازنویسی شدند. دو تابع برازندگی مختلف در نظر گرفته شده است. تابع اول یک تابع کلاسیک در این مسئله است که روی کاهش فاصله تمرکز دارد و تابع دوم توازن بار را هم در نظر می‌گیرد. نتایج پیاده‌سازی الگوریتم و اجرای آن روی نقشهٔ منطقه 1 تهران نشان می‌دهد که در صورت وجود نه مرکز امداد میانگین فاصلهٔ هر نقطه تا مرکز 760 متر می‌باشد که نسبت به بهترین مقالات قبلی حدود دو درصد بهبود را نشان می‌دهد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Location Allocation of Earthquake Relief Centers in Yazd City Based on Whale Optimization Algorithm

A. Zarepor Ashkezari[a], H. Mosalman Yazdi*[b]

[a] Department of Civil Engineering, Yazd Branch, Islamic Azad University, Yazd, Iran
[b] Department of Civil Engineering, Maybod Branch, Islamic Azad University, Maybod, Iran

PAPER INFO

Paper history:
Received 14 December 2020
Received in revised form 31 December 2020
Accepted 16 February 2021

Keywords:
Disaster Management
Earthquake
Heuristic Algorithms
Relief Center
Whale Optimization Algorithm

ABSTRACT

Despite the fact that many governments try to set rules that guarantee having resistant buildings, there are many vulnerable structures in the world. Hence, establishing earthquake relief centers is an important issue in order to control the effect of an earthquake. Iran is a country in middle east which is severely vulnerable against earthquake. Yazd is a central city in Iran. Since there is no such a study for Yazd city, this city is considered in this study. The parcels' layer of the GIS map of Yazd city has been used as the input of the problem. Since the location allocation of relief centers is a problem with huge complexity and cannot be solved in polynomial time, Whale Optimization Algorithm (WOA) has been used to solve the problem. The Whale Optimization Algorithm or The WOA is a particle based heuristic algorithm which is suitable for solving hard problems. The main contributions of the research are modifying WOA function for the problem and designing a new method for creating whales. In order to reduce the time of reaching to the reasonable solution an innovative whale generating method has been designed. The results show that average distance of each parcel from its relief center is 1541 meters and the standard deviation of 114.

doi: *10.5829/ije.2021.34.04a.04*

1. INTRODUCTION

In order to mitigate the effects of natural disasters, governments and authorities must make plans and decisions in advance [1, 2]. Natural disasters, such as earthquake and flood, are inherently unpredictable and catastrophic. Lack of comprehensive plans along with low precautionary measures can have disastrous consequences for the region. Some of the serious consequences would be loss of properties, death tolls and injuries, contagious diseases and homeless people. Reducing these harsh events is the main mission in each relief and response process [3, 4]. The stronger logistic plans you have, the easier you can fulfill these goals [5]. In other words, the main part of a rescue mission is its logistic plans. At the time of any disaster, the level of our success is at the heart of decisions that we have already made [6, 7]. Moreover, when there is a detailed plan for logistics in advance, it would be easier to coordinate communication and commuting process with the delivery of commodities [8, 9]. As a result, we will have a better response time. With increasing the speed of delivery, we will be closer to our goals in rescue mission [10, 11]. Pre-locating of the relief centers is one of the most important initiatives in order to reduce the delivery time [12, 13]. This approach was used in WWII as a military strategy in order to increase the possibility of victory and also reduce the number of wounded soldiers [14]. The whole point of logistic strategy is proper location allocation of relief centers. Location of relief centers has a profound effect on rescuing process. There are many parcels on the map of a region which need to be rescued [15, 16]. If we have a closer relief center to each parcel we will deliver the food and medicine and other necessities in more reasonable time. Out of proportion distance between the parcels and the centers would decrease the quality of helping process. Helping injured people, delivering food and making a shelter for people in need would be

*Corresponding Author Email: *HAMosalmanYazdi@gmail.com* (H. Mosalman Yazdi)

Please cite this article as: A. Zarepor Ashkezari, H. Mosalman Yazdi, Location Allocation of Earthquake Relief Centers in Yazd City Based on Whale Optimization Algorithm, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 782-792

accelerated if they are in short distance from the relief centers. Therefore, the ultimate vision of the decision makers is the relief centers to be in all people's fingertip [17, 18]. The size of the region which is affected by a natural disaster has direct effect on the complexity of the problem. When it comes to a metropolis there are many bottlenecks and congested areas on the map which can complicate the rescue process. Therefore, big cities with huge populations are exposed to more hazards and in order to decrease the risk we should have a more accurate plan and strategy [19, 20].

The logistic strategy in Yazd city in center of Iran has been addressed in this paper. This big city, like many parts of Iran, is vulnerable to earthquakes and should have an accurate plan to mitigate the effects of a possible disaster. Finding the best location for each relief center is the main goal of this research. Through decreasing the distance between centers and parcels and also balancing the load of work on each center, this research is aiming to present a solution for delivering the best service in shortest time for each person in need in Yazd city.

There is a classic problem named location allocation problem or LA which is referred to any kind of problem involved in finding suitable locations for bunch of entities. It is proved that LA is a NP-hard problem which does not have any polynomial solution. In a situation that we have a large number of objects, it is almost impossible to find the optimum solution in a reasonable time. In other words, it is true to say that solving this problem in general is not plausible [21, 22]. The most effective way to tackle this problem is using optimization methods. In this method we try to come up with an algorithm that can find a sensible solution which in close enough to the best solution to be accepted by the users. Despite the fact that these methods cannot find the optimal solution, they are useful because of their reasonable execution time. In fact, we sacrifice the best answer in order to reach a useful solution in a relatively short time [23]. One of the most famous optimization methods are particle based algorithms. These algorithms try to solve the problem by simulation of particle movements which has drawn the inspiration from natural phenomena [24, 25]. Recently a new particle based algorithm has been designed which optimize the problems by inspiration from whale's hunting process. This innovative algorithm which is called WOA (or Whale Optimization Algorithm) has added new characteristics to previous algorithms which seems to be promising to solve different types of NP-hard problems [26, 27].

There is no a comprehensive research about Location Allocation problem for Yazd city in center of Iran. Therefore, in this research a new heuristic algorithm based method has been proposed in order to tackle the problem. The WOA has been adjusted to solve the LA problem in Yazd city of Iran. The proposed method has been executed on GIS information of Yazd map in order to locate relief centers in the best places. The main aim of the paper is finding the best locations for earthquake relief centers by redesigning WOA.

To analyze the algorithm, all parcels are pulled out from a GIS map of Yazd city. In proposed algorithm, all of the locations in Yazd city could convert to a relief center. When we have some pre-defined candidate locations, the solution space will become smaller and solving the problem would be easier. The WOA has been redesigned and tuned to solve the Location Allocation problem in Yazd city. The proposed method has some parameters which are calibrated by means of a simple map. After that, the problem in Yazd city has been solved by means of proposed algorithm.

The rest of the paper is structured as follows. The second section reviews previous studies in this problem. In the third section, the proposed method has been presented completely and all of its phases have been clarified. In the fourth section, the proposed algorithm has been simulated and assessed thoroughly. In the fifth section, there is a summary, conclusion and some suggestions for future works.

2. PREVIOUS STUDIES

The Location Allocation problem or LA is a general problem which is involved in many aspects of human life. To some extent, improving the quality of life lays on solving different problems which can be boiled down to a LA problem. Therefore, in this section some solutions for various LA problems have been reviewed.

2. 1. Mathematical Model Approaches

One of the most famous and effective approaches to solve the location allocation problem is designing a mathematical model for the environment which the LA problem is supposed to be solved for it. All the parameters and their constraints can be embedded in the model. In order to reach the solution this model can be solved and optimized by different methods. Some of the recent papers which used this approach will be reviewed following.

Rahmani [28] proposed a new method for blood supply chain management. Their method is an accurate algorithm which tries to reduce the delivery time and avoid shortages in blood. The approach is locating blood supply centers in appropriate point that causes delivery process to be conducted in time. Therefore, the objective function is minimizing the distance between blood centers and people in need. The algorithm behind the model is Lagrangian Relaxation. The results show that the cost of the system is in direct proportion with the level of demand. Salehi et al. [29] designed another method for blood delivery. In this paper authors considered the situation of after earthquake. They presented a sophisticated model which takes to account some important issues about the blood, such as compatibility of the blood group. Two important variables can be optimized in this model. First the location of each relief center and second the amount of blood which should be saved in each center. The simulation results for Tehran city prove the effectiveness of the proposed model. Fazli-Khalaf et al. [30] have presented a new model for relief management. This model has three different objective functions. Many involved parts in after earthquake rescue mission have been attached to the model. This research was aiming to find the best locations for permanent and temporary relief centers in order to reduce the delivery delay. The simulation results on Tehran city showed the positive effects of the proposed model.

Boonmee et al. [31] solved the LA problem for relief facilities before and after a disaster. In order to evaluate the method, different conditions have been tested (deterministic, stochastic, dynamic, robust). Before the disaster, location allocation of shelters and warehouses was carried out and after that, location allocation of healthcare and distribution centers was focused. The proposed method has been tested in all conditions through a case study. Mahootchi and Golmohammadi [32] designed a stochastic mathematical model which has two levels. This model is executed in two phases before and after disaster. The location allocation of relief centers and assigning each parcel to a single center are conducted in the first and second phases, respectively. The level of storage in each center and the total amount of goods are also specified. The simulation results showed that the cost of relief mission and the number of needed centers are in the direct proportion to intensity of the earthquake. The proposed method was tested on Tehran city in Iran. Sebatli et al. [33] proposed a new method which can allocate necessary goods to areas in need. Their algorithm is based on a mathematical model which has two phases. Their aim was reducing the distance and the cost of delivering necessary supplies. The Yildirim region of Turkey was used in order to evaluate and validate the model. Chu et al. [34] proposed an integer nonlinear programming in order to optimize the process of assigning medical and healthcare teams to each group in the rescue mission. Their innovation was taking to account the stochastic transition probability of triage levels. In order to find a solution, the stochastic Markov chain was used. Their main goal was increasing the number of wounded which have received emergency medical care. The results show that assigning the medical teams to the closest and worst affected areas causes more lives to be saved.

2. 2. Heuristic Approaches The papers which are reviewed so far were all based on mathematical models. There is another approach to address the location allocation problem. Since the LA problem is NP-hard one, the optimization algorithms which are based on heuristic methods are very useful to find a reasonable solution in a sensible time. The particle based algorithms and the genetic based algorithms are the famous heuristic methods which have been widely used to tackle the LA problem. Here some of the recent papers with this approach will be reviewed.

Golabi et al. [35] tried to find the optimal location of distribution centers in the large-scale disasters. They assumed that it was possible there would be a problem to reach to the intended points. For optimization phase they adjusted three different heuristic algorithms to be applied to the model. The GA or genetic algorithm was the first one. Another algorithm they used was Simulated annealing, and the third one was Memetic algorithm. They tested their method through a case study in Tehran. In another paper, authors designed a stochastic mathematical model in order to minimize the effects of natural disasters. Their new algorithm, which was designed for pre-disaster time, was capable of finding the best location for each center and finding the optimal capacity for each center. They also designed an objective function in order to minimize the rate of causality. They adjusted a particle based algorithm to optimize the solution [36]. In another research authors put forward a new method which tried to compare the effectiveness of the Genetic Algorithm (GA) and Bees Algorithm (BA) in location allocation of relief centers and assigning the parcels to them. The GIS data was used as the input of the algorithm. Then, each algorithm applied to the data separately for finding nine center location between the candidates and assigning the parcels to them. Before the main testing phase the algorithm was calibrated with a simple synthetic data. The simulation results indicated that the convergence of the BA was gradual to some extent, while the behavior GA was step by step. In terms of stability both algorithms were acceptable [37]. Saeidian et al. proposed a method in which the location allocation of relief centers is specified. They used Geospatial Information System (GIS), the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) decision model, a simple clustering method and the two meta-heuristic algorithms of Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO). The authors compared PSO and ACO in different conditions. The simulation results indicate the efficiency of TOPSIS, the clustering method, and the particle based algorithms [38]. In another paper, the capacitated location-allocation problem with stochastic demand was addressed. They designed a mathematical model to find the best locations of the relief centers. A local search method was combined with genetic algorithm to tackle the problem [39]. In another research, various parameters including demand and flow of relief commodities, capacity of centers, transportation of injured people, capacity of vehicles for commodities and injured people, and back up centers were taken into account in different parts of planning. A real data from Tehran city in Iran was used to test the method. The modified multiple-objective particle swarm optimization (MMOPSO) and Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) were the heuristic algorithms which were adjusted for the problem. The simulation results show that the MMOPSO has the best effectiveness [40].

Ghasemi et al. proposed a robust simulation-optimization method for planning before the disaster. Moreover, the amount of needed goods was one of the important parameters in this method. One of the stochastic parameters of the model was the demand pattern in the problem. Robust optimization approach was used to handle uncertainty. The proposed model could specify the location of each relief center and its parcels. The genetic algorithm has been modified and used in order to optimize the solution [41].

A multi-level facility location problem (FLP) is formulated to find the best number of relief centers and their locations in literature [42]. The authors assumed that the demand pattern was based on Poisson distribution. All demands are satisfied by the closest relief center. A hybrid genetic algorithm is developed to optimize the proposed model. The effectiveness of the proposed method is tested by means of a case study. The simulation results show that there was an increase in effectiveness of relief centers and also a significant decrease in response time [42]. Table 1. summorizes some of the most important reviewed papers.

3. PROPOSED METHOD

In this article a new method has been proposed based on Whale Optimization Algorithm (WOA) in order to optimize earthquake relief centers' location in Yazd city. Since the WOA is suitable for continuous problems, it has been adjusted for solving location allocation problem. All functions have been designed from scratch. Moreover, an innovative method has been proposed to speed up the convergence of the evolutionary algorithm. In this method first generation of whales is produced in a way that is more likely to be near the optimal solution. If the first generation whales are close to the best solution, the convergence of the algorithm will be faster. Figure 1 demonstrates the steps of the proposed method.

According to Figure 1 the input of the algorithm is the polygon data, extracted from GIS of Yazd city. At the beginning of the algorithm, first random solutions are produced. As it mentioned above, first whales creating is taken place by an innovative function. After this phase, there are some functions to move the whales. After each movement, the algorithm must update the fitness values. This steps will continue until we reach to exit criteria. Reporting optimal solution is the last phase of the proposed algorithm. In following, each step of the algorithm has been explained in detail.

3. 1. Whales' Structure

Each whale represents a random solution for the problem. In other word in each solution, location of relief centers in Yazd city can be found. Random whales are not optimal but they are supposed to be improved by moving functions. Every single whale is shown by an array. The length of the array is equal to number of relief centers. Each index of array represents a particular relief center and the value inside this index demonstrates the location of this relief center. Each location is determined by a couple of values (x and y). Table 2 shows an example of a whale that includes six relief centers.

TABLE 1. Previos works

Ref.	Problem	Method	Pros & cons
[28]	Blood supply	Math model	-Reasonable solution -Low scalabiity
[29]	Blood supply	Math model	-Low running time -Low scalability
[30]	Relief Management	Math model	-Multi objective -Complicated model
[31]	Location Allocation	Math model	-Optimal solution -Low scalability
[32]	Location Allocation	Heuristic	-Scalable -Computational overhead
[37]	Location Allocation	Heuristic	-Using GIS data -Computational overhead
[38]	Location Allocation	Heuristic	-Clustering -Using GIS data
[40]	Relief management	Heuristic	-Multi objective -Using GIS data

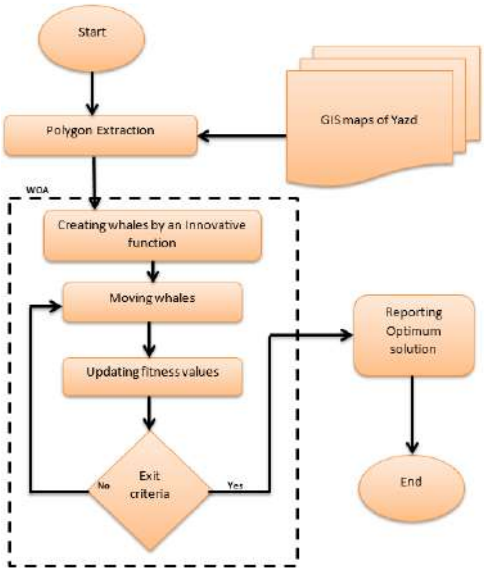


Figure 1. Steps in proposed algorithm

TABLE 2. A whale with 6 centers (meter)

Whale	$Center_1$	$Center_2$	$Center_3$	$Center_4$	$Center_5$	$Center_6$
(x , y)	(21,45)	(83,79)	(18,31)	(50,65)	(13,90)	(37,40)

3. 2. Creating Whales

In order to achieve better solution and as an aim to speed up the algorithm, initial whales are produced by an innovative function. This function has been designed in a way that tries to spread the first location of centers uniformly. Relief centers should be able to serve all parts of the region and this service should be fair. So if we spread the centers among the region it is more likely to have uniform services, however population distribution and more congested points of the map should be considered. For achieving this goal, a partitioning function has been proposed in Algorithm 1.

```
Algorithm1 : partition
Input : points of the map as (x,y)
output : lines
lineNum = ⌊√n⌋
pointNum = length(points)/lineNum;
Vline = 1;
for i = 1 to max(x)
        if (number of points with x ≥ Vline * pointNum) or (i = max(x))
                add a vertical line in this i position to lines;
                Vline = Vline + 1;
Hline = 1;
for j = 1 to max(y)
        if (number of points with y ≥ Hline * pointNum) or (i = max(y))
                add a horizontal line in this j position to lines;
                Hline = Hline + 1;
return lines
```

In Algorithm 1 *points* is an array that contains all points of the map. *LineNum* is a variable that determines the number of vertical and horizontal lines to partition the map. The number of vertical and horizontal lines are qual. *pointNum* is the number of points between each two lines. *Length(points)* is the number of points in the map. The output of the algorithm is an array, which is called *lines*, that contains all vertical and horizontal lines. The map could easily be partitioned by means of these lines. The partitioned map contains many rectangles that each one shows a particular partition of the map. Figure 2 shows output of this algorithm for two different maps. Comparing these two solutions shows that the proposed function not only considers the area of the map but brings the distribution of the points to play. In other word this algorithm tries to partition the map in a way that each part contains roughly equal number of points.

After partitioning it is time to create whales. In this step, for each whale we should select candidate points from different rectangles. Selecting candidate points from different rectangles helps us to distribute the centers among the map and make the random whales close to best solution. when number of rectangles are less than number of centers it is not a big problem to select multiple points from a single rectangle. Anyway, the number of selected points should be equal to number of centers.

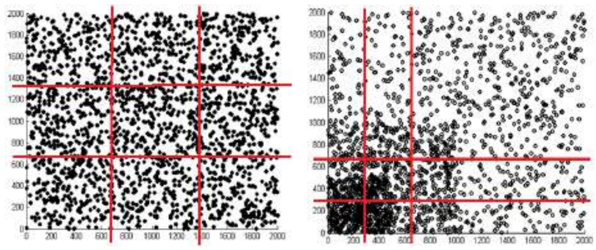
3. 3. Fitness Function

A new fitness function has been designed in order to achieve the best locations of relief centers in Yazd city. Two important issues have taken into account in calculating fitness value. Since the fitness function represents the objectives of the designers, this article has concentrated on two different aspects of the problem that are important for decision makers. The first aspect is they want to reduce the distance between centers and point as much as possible. And the second aspect is they want centers not to be over loaded. For the first goal, average distance between centers and points has been considered. The fitness function tries to reduce this value. For the second goal a penalty function has been designed. In the penalty function, finding a center that is responsible for more than average induces a negative effect on fitness value. Equation (1) shows the penalty function. The total penalty for a whale is equal the sum of the center penalties in the whale.

$$penalty(center_i) = \begin{cases} 1 & parcels_i > \dfrac{n}{k} \\ 0 & O.W. \end{cases} \tag{1}$$

Equation (2) shows the fitness function. The penalty value is in the denominator in order to have negative effect on the fitness value. The fitness function creates a value that represents the closeness of centers to the points and balancing of the centers' load simultaneously.

$$fitness(whale) = \frac{\sum_{i=1}^{n} distance(parcel_i, center_i)}{\sum_{j=1}^{k} penalty(center_j)} \tag{2}$$



a. Even map b. Uneven map

Figure 2. Partitioning example (meter)

3. 4. Move Functions The WOA has two functions to move the whales. These functions are Shrinking and Spiral respectively. Classic move functions of WOA are not useful in this problem, hence they should be redefined to meet constraints of the problem. Both of move functions use a particular concept which is called Distance. Distance value shows the difference between two given whale. In this article Distance of two whales is in the form of an array that contains the distance between nearest centers in two whales. Therefore, the length of Distance array is equal to Whale array. In following the detail of each move function has been presented.

3. 4. 1. Spiral The Spiral function tries to go toward the best answer (prey) through a spiral path. In other word in this function each whale tries to go near the best whale conservatively. So the whale spirals around the prey and approaching it. Figure 3 illustrates the rotation angle and moving toward the best answer. W wants to spiral around the Best. The rotation angle is $2\pi l$ where l is a random value in $[-1\ +1]$. The imaginary line between W and Best must be calculated by Distance function.

Algorithm 2 shows the details of Spiral. The direction of rotation is clockwise. The inputs of the function are W and Best. W is the whale that is supposed to be spiraled and Best is the prey location (best answer).

Algorithm2: Spiral

Input : $\boldsymbol{w_1}$: *a whale for spiral* , ***best*** : *the best whale*
Output : $\boldsymbol{w}$: *spiraled whale*

$l = a\ random\ number\ in\ [-1,+1];$
$\theta = 2\pi l;$ *// rotation angle*
for $i = 1$ *to* $len(w)$
    $point = a\ random\ location\ on\ line(w_1, best);$
    $destination = \theta$ *sized roundClock rotation centered by best*;
    $w(i) = destination;$
end

3. 4. 2. Shrinking The second move function is Shrinking. In Shrinking a random whale moves directly toward the prey (best answer). In this function, there is not any rotation, so in comparison with Spiral, the whales go faster toward the prey. Algorithm 3 shows the pseudo code of the Shrinking. The inputs of the function are W and Best. W is a random whale and Best is the prey.

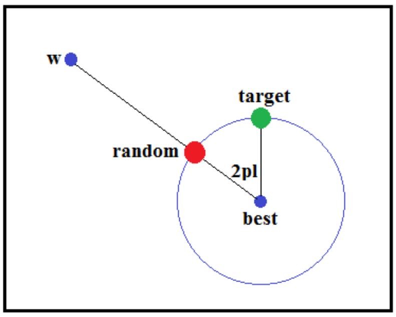


Figure 3. An example of Spiral

Algorithm3: Shrinking

Input : $\boldsymbol{w_1}$: *a whale for spiral* , ***best*** : *the best whale*
Output : $\boldsymbol{w}$: *shrinked whale*

for $i = 1$ *to* $len(w)$
    $destination = a\ random\ location\ on\ line(w_1, best);$
    $w(i) = destination;$
end

3. 4. 3. Search Apart from Shrinking and Spiral there is another useful function in proposed method which is called Search. The search function consist of three different functions. One of them is shrinking function that is used with different parameters. Instead of moving a whale toward the best, here, the shrinking function tries to move a whale toward another whale which is selected randomly. Join and random walk are two another function that are used in the search. These two functions in some extent act like crossover and mutation operators in Genetic Algorithm. All in all, the search function is used for exploring problem space. Algorithm 4 shows the pseudo code of this function.

Algorithm4 : search

Input : $\boldsymbol{w_1}$: *a whale for shrinking* , $\boldsymbol{w_2}$: *a randomly selected whale*
Output : $\boldsymbol{w}$: *a new whale*

$distance = whaleDistance(w_1, w_2);$
$tmp = a\ random\ number\ in\ [0,1];$
if $tmp < 0.25$
    $w = shrinking(w_1, w_2);$
elseif $tmp < 0.25$
    $w = join(w_1, w_2);$
else
    $w = randomwalk(w_1);$
end

3. 5. Contributions Given the details of the proposed method which has been explained in this section, the following contributions can be considered as the novelties of this work:

- Using the WOA for the first time to solve the LA problem in Yazd city.
- Modifying the WOA functions in order to solve a discrete problem.
- Introducing an innovative Whale generating algorithm which can improve the method's efficiency.

4. EXPERIMENTAL RESULTS

After presenting the proposed method, it is time to implement and evaluate it. In order to analyze the effect of the algorithm on the Yazd maps, it has been implemented in Matlab software. The first step is calculating best parameters for the method and after that we can use the algorithm for solving the problem in Yazd city maps.

4. 1. Best Parameters of the Algorithm The proposed method has some parameters which determine its performance. In order to achieve better solutions, these parameters have to be calculated. Using a simple regular map is the best way to calculate these parameters. Since the optimal solution of this map is predefined, the effect of various parameters could easily be assessed. The number of whales and the number of movements are the parameters that must be calculated.

The outputs of the proposed method with different parameters have been listed in Table 3. According to the results it could be easily understood that the best parameters is the row number nine. So in the next sections all of the results are based on 40 whales and 100 movements. the pattern of parcel assignment has been shown in Figure 4 which is completely sensible and true.

To be more specific, the proposed method has been applied to a uniform map. The optimal parameters have been used to solve the problem. This hypothetical map has 2000 parcels which are distributed on the map uniformly. It is supposed to have six relief centers on the map. The size of the map is 2000 in 2000. The map has been shown in Figure 5.

In such a map, the sensible solution should be an assignment where number of parcels are equal in all centers. Figure 6 demonstrates the assignment pattern of the solution. As it could be seen the distance between centers are almost equal. The number of parcels for each center is roughly equal 330. Given the fact that the map has 2000 parcels and six centers, 330 parcel for each center is satisfactory (2000/6=333.33). In this problem, it has been assumed that relief centers can be located in each place of the map. If we have the limitation of locating centers just in pre-specified locations, there will no problem for the algorithm because we only have the problem space smaller, in other word, solving the problem will be easier.

4. 2. Yazd City In this section the proposed method has been applied to parcels of Yazd city in order to achieve the best locations of relief centers in the city. In this city there are roughly 8500 parcels. The parcels have been extracted from parcel layer of GIS maps. The GIS maps of the city has been acquired from the Ministry of Roads and Urban Development. According to optimal parameters, the algorithm has been executed with 40

TABLE 3. Results of the proposed method with different parameters (meter)

Row#	Whale#	Move#	Center1	Center2	Center3	Center4	Fitness	Time(minute)
1	20	60	512	580	669	739	21	2.5
2	20	80	643	599	572	686	19.5	4
3	20	100	602	535	711	652	19	5
4	30	60	699	605	614	582	19	6
5	30	80	615	641	610	634	18	7
6	30	100	637	620	618	625	16.5	8
7	40	60	621	625	630	624	16	5.5
8	40	80	621	625	630	624	16	7
9	40	100	625	625	625	625	12	9.5

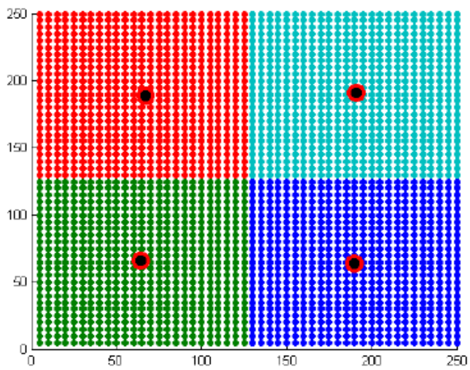


Figure 4. Optimal solution (meter)

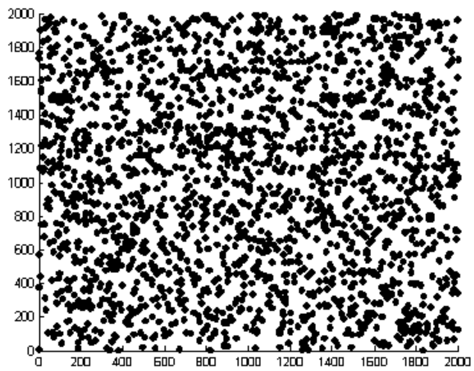


Figure 5. Sample map with 2000 parcels (meter)

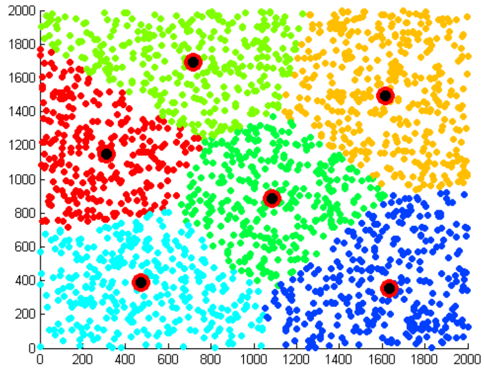


Figure 6. Location of relief centers on the map (meter)

whales and 100 movements. Figure 7 shows the results for Yazd city. In order to eliminate the mantissa, the numbers of the chart are 10000 times bigger than the real map. Figure 8 shows the process of optimization. The X axis shows the round of execution and the Y axis shows the fitness value.

4. 3. Effectiveness of Whale Creating Method

In this article a new method has been proposed in order to create effective whales at the beginning of the algorithm (section 3.2). The effectiveness of this algorithm has been analyzed in this section. For analyzing this method, map of Yazd city has been used. The best solutions, with and without whale creating method, have been compared. Three important metrics can indicate the effectiveness of the method. The first metric is convergence speed which should be shown with a chart. The second metric is average distance that shows the average distance between each parcel and its center. Finally, the third metric is standard deviation which shows that each parcel how fair could be rescued. Figure 9, demonstrates the optimization of objective function. The red line in Figure 8 shows the execution of the algorithm with random whales and blue line shows it with whales which created with proposed method. The speed of convergence and the final value in blue line in better than red line. Better solution in blue line is because of having better whales at the beginning of the algorithm. In fact, when we use the innovative whale generating method, due to having better whales at the begging of the algorithm, the final solution is slightly more optimized.

Table 4 shows the value of average distance and standard deviation for two executions. The first line shows the results for random whales and the second line shows them for whales which created by proposed method. The results show that by using whale creating method, both of metrics have more optimal values.

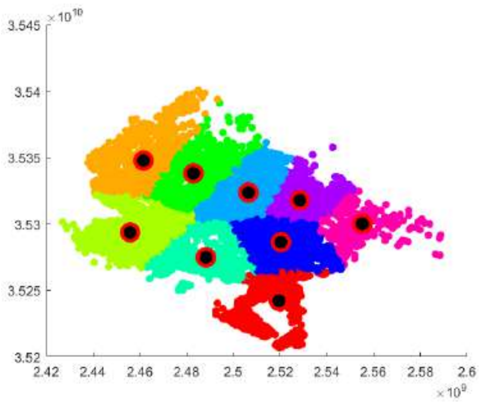


Figure 7. Location of relief centers in Yazd (meter*10000)

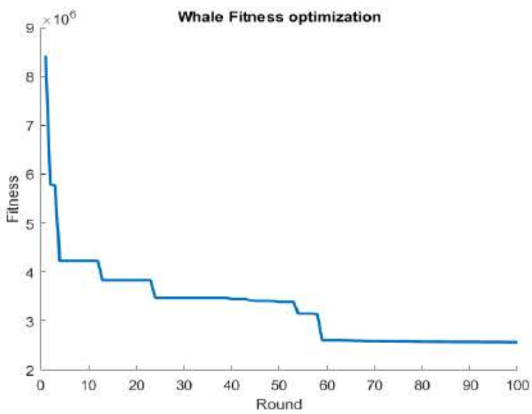


Figure 8. Optimization process for Yazd

4. 4. Stability of the Proposed Algorithm

The proposed algorithm tries to solve the problem by means of whale optimization algorithm. WOA is a kind of

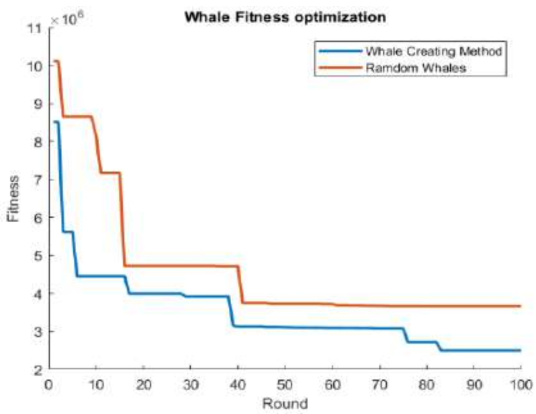
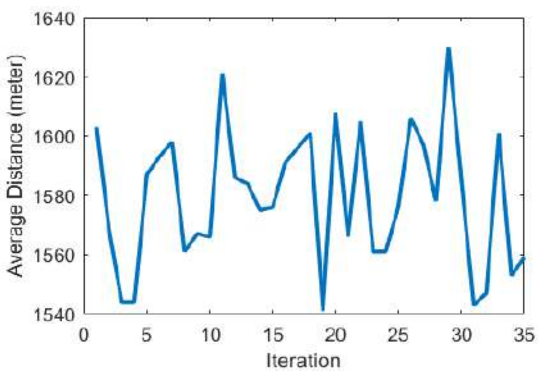


Figure 9. Effect of whale creating method

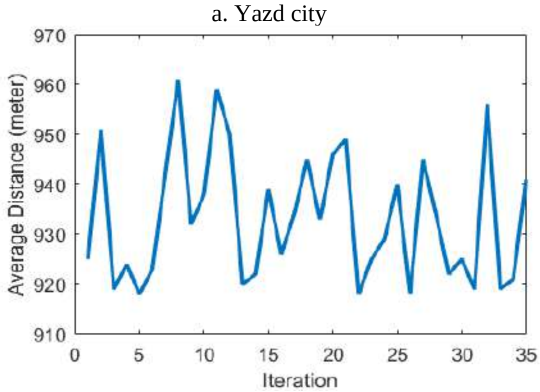
TABLE 4. Value of average distance and standard deviation for two executions

Whale creating method	Avg_dis (meter)	Standard Deviation
random	1590	155
proposed	1541	114

algorithm that solves problem heuristically. These kind of algorithms produce a different solution in each execution. Hence, the proposed method should be analyzed in terms of stability. In order to assess the stability of the algorithm it has been executed 35 times for Yazd city. The results show that the average distance is equal 1580 and the standard deviation is 20, whereas the best solution is 1541. For better assessment, the proposed algorithm has been applied to a hypothetical map as well. The results for 35 execution show that the average distance is equal 933 and the standard deviation is 13, whereas the best solution is 918. For both maps the results are satisfactory. Figure 10 shows the details of the results. Figure 10(a) is for Yazd city and Figure 10(b) is for hypothetical map.



a. Yazd city



b. Hypothetical map

Figure 10. Stability of the proposed method (by 35 executions)

5. CONCLUSION

In this paper, a new algorithm has been proposed in order to find the best locations for earthquake relief centres in Yazd city. The Whale optimization algorithm or WOA has been modified for using in this problem. Redesigning functions of WOA by means of some innovative concepts is one of the most important contributions in this article. Another contribution is introducing a new method for creating better whales at the beginning of the algorithm. By means of better whales the algorithm could be converged faster and produce better solutions. Some hypothetical maps have been used to calculate the best parameters of the algorithm. Moreover, results of the problem for Yazd city has been reported thoroughly. The best average distance for each relief centre is 1541m and the standard deviation is 114. After evaluating the stability of the algorithm it can be understood that the results of the algorithm are reliable. For future works, parameters like vulnerability of the buildings in Yazd city, Level of relief centres can be considered in designing algorithm.

6. REFERENCES

1. Bozorgi-Amiri, A., Jabalameli, M. S., Alinaghian, M., and Heydari, M., "A modified particle swarm optimization for disaster relief logistics under uncertain environment", ***The International Journal of Advanced Manufacturing Technology***, Vol. 60, Nos. 1–4, (2012), 357–371. doi:10.1007/s00170-011-3596-8
2. Beiki, H., Seyedhosseini, S. M., Ghezavati, V. R., and Seyedaliakbar, S. M., "Multi-objective optimization of multi-vehicle relief logistics considering satisfaction levels under uncertainty", ***International Journal of Engineering, Transactions B: Applications***, Vol. 33, No. 5, (2020), 814–824. doi:10.5829/IJE.2020.33.05B.13
3. Every, D., and Richardson, J., "A framework for disaster resilience education with homeless communities", ***Disaster Prevention and Management: An International Journal***, Vol. 27, No. 2, (2018), 146–158. doi:10.1108/DPM-08-2017-0196
4. Beiki, H., Seyedhosseini, S. M., Ghezavati, V. R., and Seyedaliakbar, S. M., "A location-routing model for assessment of the injured people and relief distribution under uncertainty", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 7, (2020), 1274–1284. doi:10.5829/ije.2020.33.07a.14
5. Liberatore, F., Pizarro, C., de Blas, C. S., Ortuño, M. T., and Vitoriano, B., "Uncertainty in Humanitarian Logistics for Disaster Management. A Review", ***Decision Aid Models for Disaster Management and Emergencies***, (2013), 45–74 Atlantis Press, Paris. doi:10.2991/978-94-91216-74-9_3
6. Maharjan, R., and Hanaoka, S., "A multi-actor multi-objective optimization approach for locating temporary logistics hubs during disaster response", ***Journal of Humanitarian Logistics and Supply Chain Management***, Vol. 8, No. 1, (2018), 2–21. doi:10.1108/JHLSCM-08-2017-0040
7. Kaviyani-Charati, M., Souraki, F. H., and Hajiaghaei-Keshteli, M., "A Robust Optimization Methodology for Multi-objective Location-transportation Problem in Disaster Response Phase under Uncertainty", ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 11, (2018), 1953–1961. doi:10.5829/ije.2018.31.11b.20
8. Fazelabdolabadi, B., and Golestan, M. H., "Towards Bayesian Quantification of Permeability in Micro-scale Porous Structures – The Database of Micro Networks", ***HighTech and Innovation Journal***, Vol. 1, No. 4, (2020), 148–160. doi:10.28991/HIJ-2020-01-04-02
9. Bozorgi-Amiri, A., Jabalameli, M. S., and Mirzapour Al-e-Hashem, S. M. J., "A multi-objective robust stochastic programming model for disaster relief logistics under uncertainty", ***OR Spectrum***, Vol. 35, No. 4, (2013), 905–933. doi:10.1007/s00291-011-0268-x

10. Kılcı, F., Kara, B. Y., and Bozkaya, B., "Locating temporary shelter areas after an earthquake: A case for Turkey", ***European Journal of Operational Research***, Vol. 243, No. 1, (2015), 323–332. doi:10.1016/j.ejor.2014.11.035

11. Karimi, B., Bashiri, M., and Nikzad, E., "Multi-commodity Multimodal Splittable Logistics Hub Location Problem with Stochastic Demands", ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 11, (2018), 1935–1942

12. Vahdani, B., Veysmoradi, D., Noori, F., and Mansour, F., "Two-stage multi-objective location-routing-inventory model for humanitarian logistics network design under uncertainty", ***International Journal of Disaster Risk Reduction***, Vol. 27, (2018), 290–306. doi:10.1016/j.ijdrr.2017.10.015

13. Borowski, P. F., "New Technologies and Innovative Solutions in the Development Strategies of Energy Enterprises", ***HighTech and Innovation Journal***, Vol. 1, No. 2, (2020), 39–58. doi:10.28991/HIJ-2020-01-02-01

14. Das, R., "Disaster preparedness for better response: Logistics perspectives", ***International Journal of Disaster Risk Reduction***, Vol. 31, (2018), 153–159. doi:10.1016/j.ijdrr.2018.05.005

15. Sadidi, J., Fakourirad, E., and Zeaieanfirouzabadi, P., "Designing a spatial cloud computing system for disaster (earthquake) management, a case study for Tehran", ***Applied Geomatics***, Vol. 10, No. 2, (2018), 99–111. doi:10.1007/s12518-018-0203-9

16. Haji Gholam Saryazdi, A., and Poursarrajian, D., "Qualitative System Dynamics Model for Analyzing of Behavior Patterns of SMEs", ***HighTech and Innovation Journal***, Vol. 2, No. 1, (2021), 9–19. doi:10.28991/HIJ-2021-02-01-02

17. Tavana, M., Abtahi, A.-R., Di Caprio, D., Hashemi, R., and Yousefi-Zenouz, R., "An integrated location-inventory-routing humanitarian supply chain network with pre- and post-disaster management considerations", ***Socio-Economic Planning Sciences***, Vol. 64, (2018), 21–37. doi:10.1016/j.seps.2017.12.004

18. Wedawatta, G., Kulatunga, U., Amaratunga, D., and Parvez, A., "Disaster risk reduction infrastructure requirements for South-Western Bangladesh", ***Built Environment Project and Asset Management***, Vol. 6, No. 4, (2016), 379–390. doi:10.1108/BEPAM-06-2015-0022

19. Noyan, N., and Kahvecioğlu, G., "Stochastic last mile relief network design with resource reallocation", ***OR Spectrum***, Vol. 40, No. 1, (2018), 187–231. doi:10.1007/s00291-017-0498-7

20. Baylan, E. B., "A Novel Project Risk Assessment Method Development via AHP-TOPSIS Hybrid Algorithm", ***Emerging Science Journal***, Vol. 4, No. 5, (2020), 390–410. doi:10.28991/esj-2020-01239

21. Allahbakhsh, M., Arbabi, S., Galavii, M., Daniel, F., and Benatallah, B., "Crowdsourcing planar facility location allocation problems", ***Computing***, Vol. 101, No. 3, (2019), 237–261. doi:10.1007/s00607-018-0670-1

22. Mihajlović, G., and Živković, M., "Sieving Extremely Wet Earth Mass by Means of Oscillatory Transporting Platform", ***Emerging Science Journal***, Vol. 4, No. 3, (2020), 172–182. doi:10.28991/esj-2020-01221

23. Babicz, D., Tihanyi, A., Koller, M., Rekeczky, C., and Horvath, A., "Simulation of an Analogue Circuit Solving NP-Hard Optimization Problems", In 2019 IEEE International Symposium on Circuits and Systems (ISCAS), (2019), 1–5. doi:10.1109/ISCAS.2019.8702694

24. Ozturk, B., "Preliminary seismic microzonation and seismic vulnerability assessment of existing buildings at the city of nigde, turkey", In 14th World Conference on Earthquake Engineering, Beijing, China, (2008), 1–7.

25. Ozturk, B., "Application of preliminary microzonation and seismic vulnerability assessment in a city of medium seismic risk in turkey", In 5th International Conference on Earthquake Geotechnical Engineering, Santiago, Chile, (2011), 1–11.

26. Mirjalili, S., and Lewis, A., "The Whale Optimization Algorithm", ***Advances in Engineering Software***, Vol. 95, (2016), 51–67. doi:10.1016/j.advengsoft.2016.01.008

27. Mehranfar, N., Hajiaghaei-Keshteli, M., and Fathollahi-Fard, A. M., "A Novel Hybrid Whale Optimization Algorithm to Solve a Production-Distribution Network Problem Considering Carbon Emissions", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 32, No. 12, (2019), 1781–1789. doi:10.5829/ije.2019.32.12c.11

28. Rahmani, D., "Designing a robust and dynamic network for the emergency blood supply chain with the risk of disruptions", ***Annals of Operations Research***, Vol. 283, Nos. 1–2, (2019), 613–641. doi:10.1007/s10479-018-2960-6

29. Salehi, F., Mahootchi, M., and Husseini, S. M. M., "Developing a robust stochastic model for designing a blood supply chain network in a crisis: a possible earthquake in Tehran", ***Annals of Operations Research***, Vol. 283, Nos. 1–2, (2019), 679–703. doi:10.1007/s10479-017-2533-0

30. Fazli-Khalaf, M., Khalilpourazari, S., and Mohammadi, M., "Mixed robust possibilistic flexible chance constraint optimization model for emergency blood supply chain network design", ***Annals of Operations Research***, Vol. 283, Nos. 1–2, (2019), 1079–1109. doi:10.1007/s10479-017-2729-3

31. Boonmee, C., Arimura, M., and Asada, T., "Facility location optimization model for emergency humanitarian logistics", ***International Journal of Disaster Risk Reduction***, Vol. 24, (2017), 485–498. doi:10.1016/j.ijdrr.2017.01.017

32. Mahootchi, M., and Golmohammadi, S., "Developing a new stochastic model considering bi-directional relations in a natural disaster: a possible earthquake in Tehran (the Capital of Islamic Republic of Iran)", ***Annals of Operations Research***, Vol. 269, Nos. 1–2, (2018), 439–473. doi:10.1007/s10479-017-2596-y

33. Sebatli, A., Cavdur, F., and Kose-Kucuk, M., "Determination of relief supplies demands and allocation of temporary disaster response facilities", ***Transportation Research Procedia***, Vol. 22, (2017), 245–254. doi:10.1016/j.trpro.2017.03.031

34. Chu, X., and Zhong, Q., "Post-earthquake allocation approach of medical rescue teams", ***Natural Hazards***, Vol. 79, No. 3, (2015), 1809–1824. doi:10.1007/s11069-015-1928-y

35. Golabi, M., Shavarani, S. M., and Izbirak, G., "An edge-based stochastic facility location problem in UAV-supported humanitarian relief logistics: a case study of Tehran earthquake", ***Natural Hazards***, Vol. 87, No. 3, (2017), 1545–1565. doi:10.1007/s11069-017-2832-4

36. Paul, J. A., and MacDonald, L., "Location and capacity allocations decisions to mitigate the impacts of unexpected disasters", ***European Journal of Operational Research***, Vol. 251, No. 1, (2016), 252–263. doi:10.1016/j.ejor.2015.10.028

37. Saeidian, B., Mesgari, M. S., and Ghodousi, M., "Evaluation and comparison of Genetic Algorithm and Bees Algorithm for location–allocation of earthquake relief centers", ***International Journal of Disaster Risk Reduction***, Vol. 15, (2016), 94–107. doi:10.1016/j.ijdrr.2016.01.002

38. Saeidian, B., Mesgari, M., Pradhan, B., and Ghodousi, M., "Optimized Location-Allocation of Earthquake Relief Centers Using PSO and ACO, Complemented by GIS, Clustering, and TOPSIS", ***ISPRS International Journal of Geo-Information***, Vol. 7, No. 292, (2018), 1–25. doi:10.3390/ijgi7080292

39. Thumronglaohapun, S., "Heuristic Methods for the Capacitated Location- Allocation Problem with Stochastic Demand", ***International Journal of Industrial and Systems Engineering***, Vol. 14, No. 6, (2020), 452–457

40. Ghasemi, P., Khalili-Damghani, K., Hafezalkotob, A., and Raissi,

S., "Uncertain multi-objective multi-commodity multi-period multi-vehicle location-allocation model for earthquake evacuation planning", ***Applied Mathematics and Computation***, Vol. 350, (2019), 105–132. doi:10.1016/j.amc.2018.12.061

41. Ghasemi, P., and Khalili-Damghani, K., "A robust simulation-optimization approach for pre-disaster multi-period location–allocation–inventory planning", ***Mathematics and Computers in Simulation***, Vol. 179, (2021), 69–95. doi:10.1016/j.matcom.2020.07.022

42. Shavarani, S. M., "Multi-level facility location-allocation problem for post-disaster humanitarian relief distribution", ***Journal of Humanitarian Logistics and Supply Chain Management***, Vol. 9, No. 1, (2019), 70–81. doi:10.1108/JHLSCM-05-2018-0036

Persian Abstract

چکیده

علیرغم اینکه دولت‌ها قوانین سختگیرانه¬ای را برای تضمین استحکام ساختمان¬ها وضع می¬کنند؛ همچنان ساختمان¬های زیادی وجود دارد که در برابر زلزله آسیب¬پذیر است. ازاین رو ایجاد مراکز امداد زلزله جهت کنترل اثرات آن بسیار ضروری است. ایران یکی از کشورهای خاورمیانه است که آسیب¬پذیری زیادی در برابر زلزله دارد. شهر یزد در مرکز ایران قرار گرفته است. از آنجایی که چنین تحقیقی برای شهر یزد وجود ندارد، این شهر در این تحقیق مورد بررسی قرار گرفته است. لایه پارسل از نقشه جی آی اس شهر یزد بعنوان ورودی مساله در نظر گرفته شده است. از آنجایی که مکان¬یابی مراکز زلزله یک مساله پیچیده بوده و در زمان چندجمله¬ای قابل حل نیست؛ از الگوریتم اکتشافی بهینه¬سازی نهنگ برای حل مساله استفاده شده است. الگوریتم بهینه¬سازی نهنگ یک روش مبتنی بر ذره است که کاربرد زیادی در حل مسائل سخت دارد. نوآوری اصلی این مقاله تنظیم توابع نهنگ برای استفاده در مساله مکان¬یابی و همچنین معرفی روش جدیدی برای ایجاد نهنگ است. جهت رسیدن سریع‌تر به جواب منطقی یک روش مبتکرانه برای تولید نهنگ طراحی شده است. بررسی نتایج الگوریتم نشان می¬دهد که میانگین فاصله هر ذره تا نزدیکترین مرکز 1850 متر و انحراف از معیار اجراهای مختلف نیز برابر 20 است.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 793-802

International Journal of Engineering

Journal Homepage: www.ije.ir

Numerical Study on Flexural Behavior of Concrete Beams Strengthened with Fiber Reinforced Cementitious Matrix Considering Different Concrete Compressive Strength and Steel Reinforcement Ratio

Z. R. Aljazaeri*, Z. Al-Jaberi

Civil Engineering Department, Al-Nahrain University, Baghdad Governorate, Iraq

PAPER INFO

Paper history:
Received 16 October 2020
Received in revised form 05 February 2021
Accepted 08 February 2021

Keywords:
Cohesive Bond Models
Fiber Reinforced Cementitious Matrix
Finite Element
Numerical Study
Reinforced Concrete Beams Strengthening

ABSTRACT

Concrete structures retrofitted with fiber reinforced cementitious matrix (FRCM) have become widespread due to their mechanical and durability performances. However, the behavior of FRCM - strengthened RC members under service loads is still a concern, and more efforts need to be done. In this study, a nonlinear three-dimensional finite element (FE) model has been developed to study the performance of reinforced concrete (RC) beams strengthened by (FRCM). The model was validated against the experimental results gathered from six beams tested under three-points bending. Consequently, the primary numerically studied parameters were longitudinal steel reinforcement ratio and concrete compressive strength. A cohesive damage parameters were investigated to represent the experimental results. Also, the theoretical flexural capacity of strengthened beams based on ACI-549.4R code was evaluated based on the numerical method results. As a conclusion, the numerical results are in a very good agreement with the experimental ones regarding yielding load, ultimate load, and failure mode. In addition, the developed models from parametric studies concluded the insignificant effect of concrete compressive strength on increasing the ultimate capacity of strengthened beam. However, the steel reinforcement ratio has a major impact on enhancing the ultimate capacity of strengthened beams.

***doi**: 10.5829/ije.2021.34.04a.05*

NOMENCLATURE

E_c	Concrete modulu of elasticity	M_n	Nominal flexural strength at section
f_c'	Specified compressive strength of concret	M_s	Steel flexural strength at section
f_{ct}	Conceret tensile strength	σ_c	Extreme fiber concret compressive stress
M_f	Fiber flexural strength at section	ε_c	Extreme fiber concret compressive strain

1. INTRODUCTION

During the last decade, FRCM composite material was developed with almost the same advantages of FRP strengthening technique such as high strength to weight ratio, corrosion resistant, and ease of implementation in addition to that to overcome some of the FRP drawbacks specially those related to fire resistance or installing on wet surfaces issue [1-2]. The other physical benefits of FRCM strengthening system are good reversibility and good vapor permeability in addition to consider the matrices not toxic material like the epoxy that utlized in FRPs technique [3-5].

On behalf of that, many experimental studies have been investigated the structural and durability performances of FRCM material as a strengthening or a repairing system for infrastructural members and compared with structures strengthened with FRP. The experimental works concluded the effectiveness of the FRCM material in increasing the ultimate flexural or shear loads of reinforced concrete (RC) beams/slabs and masonry walls (references of different aspects)[6-10]. Also, the FRCM material was used to improve the

*Corresponding Author Institutional Email: *zracnb@mst.edu* (Z. R. Aljazaeri)

Please cite this article as: Z. R. Aljazaeri, Z. Al-Jaberi, Numerical Study on Flexural Behavior of Concrete Beams Strengthened with Fiber Reinforced Cementitious Matrix Considering Different Concrete Compressive Strength and Steel Reinforcement Ratio, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 793-802

confinement of RC columns, repairing corroded RC beams, and seismic upgrading for other structural elements [11-13].

Most of the previous studies focused on theoritical and numerical studies of structures strengthened with FRP systems [14-17]. A numerical finite element analysis of RC beams strengthened with EB-FRP techniques was implemented [18]. According to this study, the validated model was able to predict maximum load capacity, load-deflection curves, and the bond slib behavior for both interfaces (FRP-epoxy resin and the epoxy resin with substrate. Furthermore, parametric study was conducted to evaluate many factors and their impact on the efficiency of strengthened RC beams. Recent advances in the use of FRCM material has highlighted the need to carry out reliable numerical analyses based on FE models to predict the strengthening or repairing behavior of FRCM composite with different variables. However, a few FE studies have been modeled structural elements strengthened with FRCM material. To the authors' best knowledge, the studies by D'Ambrisi [4] and Ombres [19] were the only ones found in the literature on this topic. D'Ambrisi et al. [4] conducted a nonlinear FE model for RC beams strengthened with FRCM composite under flexural loads. A perfect bond between the PBO-fibers and concrete was assumed. The numerical load-deflection curves were fit the experimental ones up to a certain point then the theoretical curves began to diverge. So the adopted models did not present the FRCM debonding failure mechanism. Ombres et al. [19] also modeled FRCM strengthened RC beams in flexural by assuming a perfect bond between the FRCM composite and concrete substrate. The numerical results were provided non-accurate predictions for load-deflection relations concerning to the experimental ones. As well, both the ultimate capacity and debonding strains of the numerical model differed in a range of 3 to 40% from that of the experimental work. On the other hand, several FE studies were conducted a single-lap direct shear test to characterize the bond-slip relation between FRCM composite and concrete substrate. D'Antino et al. [20] analyzed the axial strain profiles of the FRCM composite-concrete interface to investigate the stress-transfer mechanism at the matrix–fiber interface. The ultimate debonding load was predicted using the expected fracture parameters from the strain profiles of seven tested specimens. The ultimate loads were in good agreement with the experimental applied loads. Focacci et al. [21] used an indirect method to define interfacial shear stress-slip relation between FRCM composite and concrete substrate. Different cohesive material laws were adopted and calibrated to simulate successfully the experimental load responses. Carloni et al. [22] used a fracture mechanic approach to account for the mixed failure modes observed in the single-lap shear test. The FE outcomes presented the experimental load responses correctly and imitated different experimental failure modes. D'Antino et al. [23] proposed an analytical model to define the bond behavior of the FRCM-concrete interface using trilinear cohesive material laws based on the experimental results of the single-lap shear test. The proposed model revealed the load response of the FRCM-concrete interface up to the peak stress with sufficient accuracy. However, the proposed model did not present the post-peak behavior of the FRCM-concrete interface. Zou et al. [24] defined a new bond-slip equation for FRCM-concrete joints based on longitudinal fiber strains. The ultimate debonding load showed good agreement with the test results of seven single-lap shear specimens. Despite all these efforts, no full bond-slip relation was proposed and standardized for the FRCM-concrete interface. On the other hand, no numerical studies conducted to validate the use of the proposed bond-slip laws in analyzing structural members strengthened with FRCM composite. There is a lack of experimental researches on the flexural behavior of RC beams strengthened with FRCM under different concrete compressive strength or longitudinal steel reinforcement ratio. This paper aims to fill these gaps numerically by presenting the results of theses factor through validated model. Validated FE models can be very robust analytical tool, since it will typically result in enormous reductions of time and cost.

2. SIGNIFICANCE AND OBJECTIVES

In many instances, flexural concrete members sustain damage due to excessive load or harsh environmental conditions and therefore require rehabilitation and upgrade, conventionally by using advanced composite such as NSM-FRP bars or EB-FRP sheets. The second generation of advanced composite (FRCM system) has several distinct properties that provides the advantages of FRP technique in addition to eliminate the problems related to organic adhesive.

The behavior of FRCM technique in strengthening RC beams considering different types of open mesh fibers, or concrete has not been fully examined and understood. The objective of this paper is to assess the applicability of cohesive bond model (CBM) in modeling FRCM strengthened beams under flexural loads and to propose and validate a robust, non-linear (FE) model that used to predict the behavior of strengthened beams considering different parameters. The validated model is used mainly to examine key parameters that have not been studied yet, including: different compressive strength of concrete, ranging from normal to high strength and longitudinal steel reinforcement ratio, ranging from (0.24 to 1.06%). Knowing the influence of these key parameters on the structural performance of strengthened beams with FRCM composite would assist in advanced applications and provide a design clue for different cases that have not been touched experimentally. It should be noticed that the various steel reinforcement ratio was used to evaluate the accuracy of ACI 549.4R total flexural capacity equation.

3. SUMMARY OF EXPERIMENTAL PROGRAM

The RC concrete beams tested by Babaeidarabad et al. [3] were selected to validate the predictions of the FE model developed in this study. Six RC beams were tested under three- points bending. Three beams made of low-strength and the other three made of high-strength concrete. The low and high compressive strength were represented the range of compressive strength of fabricated beams in the last decades. So based on ACI 363R [25], the concrete with compressive strength of 34 MPa (5000 psi) was considered high strength. Nowadays, these beams are in need of repair and strengthen to meet the new requirement of building codes and retain back to service.

All the beams had a rectangular cross-section of 152 mm*260 mm and a length of 1,829 mm with a clear span of 1,524 mm. The six beams were reinforced with 2-13 mm diameter longitudinal steel bars at the tension and compression faces and 10 mm diameter transverse reinforcement spaced at 127 mm. Typical beam layout and corss-section is presented in Figure 1.

The specimen identification system consisted of two parts as summarized in Table 1. The first part represented concrete information, L for Low compressive strength and H for high compressive strength concrete. The second part identifies the number of open mesh layers. In case of reference specimen, the second part denoted as control.

4. FINITE ELEMENT MODEL DEVELOPMENT

4. 1. Model Description Finite element analysis was performed the nonlinear behavior of RC beams strengthened with FRCM composite using Abaqus 6.14

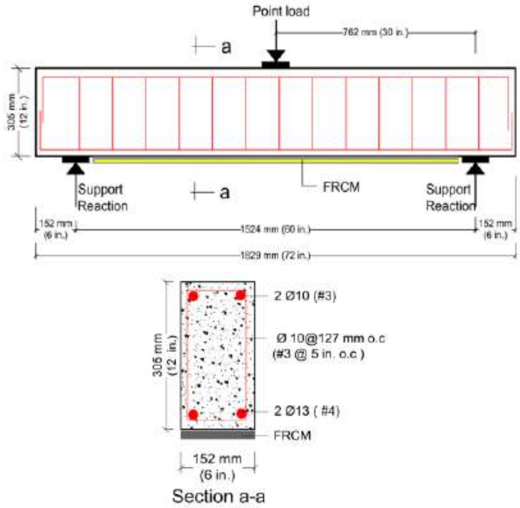


Figure 1. Beam layout and its cross-section

/standard [26]. The full length of tested beams was simulated to recognize the failure mechanism of FRCM composite along the beams. A three-dimensional finite element mesh is shown in Fig. 2. A fine mesh of approximate global size (25 mm) was selected to ensure sufficient accuracy in the numerical results. Also, the selected meshing size providing a good balance between accuracy and the cost in terms of desk space and running time. The boundary conditions were set in the middle of supporting plates. A simply-supported boundary was considered in the x and y directions, respectively. The roller support boundary was constraint in the x-direction only. The load was induced by the deflection-controlled method in order to simulate the post-peak behavior. The number of loading steps was 100 step. The total applied load was equal to the sum of two vertical reactions associated with each loading step.

4. 2. Types of Elements The beam element C3D8R was adopted for both concrete and steel plates that used under loading and supporting points. The truss element T3D2 was adopted for modeling the steel reinforcement (longitudinal and transverse reinforcement). The steel reinforcement was embedded inside the concrete beam. This type of bonding did not include the effect of slip between the reinforcement and concrete beam. Instead, these properties were partly considered through the definition of concrete tension softening. The shell element S4R was adopted for FRCM composite. The loading and supporting plates were tied to the concrete beam by surface to surface contact.

4. 3. Material Models The compressive behavior of concrete material was first characterized by a linear elastic behavior and second by a nonlinear plastic behavior (Figure 3a). The linear elastic behavior was

TABLE 1. Experimental test matrix

Specimen ID	Type of concrete	Number of layers
L-control	Low compressive strength	-
L-1ply	Low compressive strength	1
L-4ply	Low compressive strength	4
H-control	High compressive strength	-
H-1ply	High compressive strength	1
H-4ply	High compressive strength	4

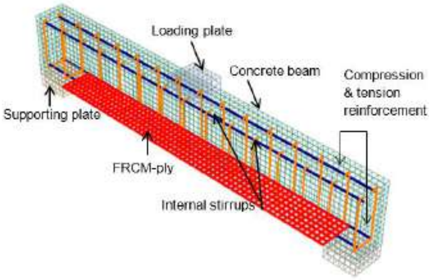


Figure 2. Typical 3D finite element mesh model for the strengthened beam

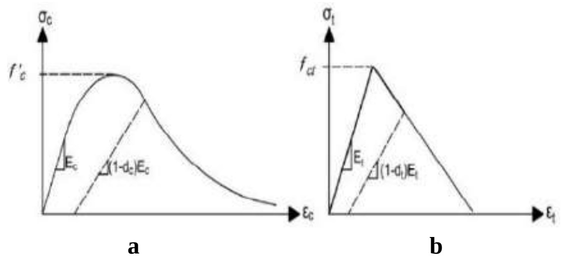


Figure 3. The behavior of concrete in uniaxial loadings: (a) compression; and (b) tension

defined by the elastic modulus (25 GPa) based on Equation (1) [27] and Poisson's ratio (0.2). A plastic damage model presented the plastic compressive behavior of concrete material. This model had two failure modes (tensile cracking and compressive crushing) [24]. The plastic damage model of concrete defined by the plastic damage parameters, density, and post-peak tensile/compressive behavior. The five damage parameters were the dilation angle (30), the flow potential eccentricity (0.1), the ratio of initial biaxial compressive yield stress to initial uniaxial compressive yield stress (1.16), the ratio of the second stress invariant on the tensile meridian to that on the compressive meridian (0.667), and the viscosity parameter (0.001). The concrete density (2400 kg/m^3) was considered. The compressive behavior of concrete was modeled by the stress-strain relation in Equation (2) [28]. The peak concrete strain was assumed to equal 0.0025 mm/mm and the ultimate concrete strain was equal to 0.003 mm/mm [3]. The tensile behavior of concrete material consisted of two phases (Figure 3b). The first phase presented the linear elastic behavior of concrete up to reaching its tensile strength. The second phase presented the descending branch in the uniaxial tensile stress-strain relation due to crack occurrence and its propagation in concrete material. The ultimate tensile strength of concrete was estimated by Equation (3) [28]. The second phase was assumed as a linear softening branch where the ultimate tensile strain at the end softening was set to be 0.001 mm/mm. The degradation in the concrete stiffness was simulated by concrete damage parameters (d_c in compression and d_t in tension)

$$E_c = 4700\sqrt{f_c'} \tag{1}$$

where f_c' is given in MPa.

$$\sigma_c = \frac{E_C\ \varepsilon_c}{1+(R+R_E-2)\left(\frac{\varepsilon_c}{\varepsilon_0}\right)-(2R-1)\left(\frac{\varepsilon_c}{\varepsilon_0}\right)^2+R\left(\frac{\varepsilon_c}{\varepsilon_0}\right)^3} \tag{2}$$

where

$R = \frac{R_E\,(R_\sigma-1)}{(R_\varepsilon-1)^2} - \frac{1}{R_\varepsilon}$, $R_E = \frac{E_C}{E_0}$, $E_0 = \frac{f_c'}{\varepsilon_0}$ and

$\varepsilon_0 = 0.0025$, $R_\varepsilon = 4$, $R_\sigma = 4$, as reported in [28]

$$f_{ct} = 0.33\sqrt{f_c'} \tag{3}$$

The steel reinforcement was simulated as an elastic-plastic material. In the linear elastic range, the behavior was defined by the young modulus (190,000 MPa) and the Poisson's ratio (0.3); whereas in the plastic range, the density (7800 kg/m^3) and the yield strength (276 MPa) were defined. The elastic-plastic properties of steel reinforcement were based on the results of coupon tests [3]. The ultimate plastic strain was assumed 0.015 mm/mm.

The FRCM composite consisted of the polyparaphenylene benzobisoxazole (PBO) fabric with a cement based curing agent. The width of PBO-fabric was 5 mm and 10 mm in the longitudinal and transverse directions, respectively. The free spacing between the strands was approximately 5 mm and 15 mm in the longitudinal and transverse directions, respectively. The two dimensional mesh of PBO fabric were embedded inside the cement based (matrix) that produced the FRCM composite. The interaction between the PBO-fabric and the cement matrix were represented by a cohesive element having initial elastic stiffness (k_0) in kN/mm, maximum shear stress (τ_{max}) $in\ MPa$, and fracture energy (G_f) in N/mm, are ranged between (1.5-9.85), (0.2-1.5), and (0.4-5); respectively.

The FRCM composite was modeled as a laminate. The FRCM composite' tensile properties were based on coupon tests reported by Babaeidarabad et al. [3]. The input parameters were: the modulus of elasticity (128,000 MPa), Poisson's ratio (0.3), and ultimate tensile strength (1,664 MPa). The FRCM laminate properties in the two directions were assumed equally. The FRCM laminate thickness was 5 mm for one-ply and 10 mm for four plies, respectively, all other properties can be found in literature [3].

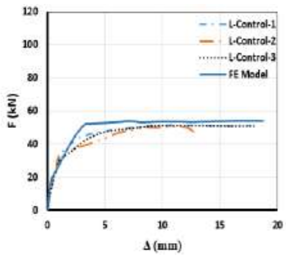
4. 4. FRCM-Concrete Interface Model The simulation of the interface between FRCM composite and concrete substrate was the critical feature in modeling the strengthened beams. Therefore, different

interfacial bond models were considered here. The cohesive bond model (CBM) was defined by a simple bilinear traction–separation law [26]. The interfacial cohesive behavior was initially linear elastic and followed by a descending linear branch which denoted the initiation of damage. The traction–separation response was described by three failure modes: opening (mode I), sliding I (mode II), and sliding II (mode III).
Through literature, many experimental and numerical works have been determined the cohesive damage parameters between the FRCM composite and concrete [13-17] and these values will guide the researcher in the future studies. Based on these studies, the CBM parameters including initial elastic stiffness (k_0) in kN/mm, maximum shear stress $(\tau_{max})\ in\ MPa$, and fracture energy $(\ G_f)\ in$ N/mm, are ranged between (1.5-9.85), (0.2-1.5), and (0.24-1.5); respectively.

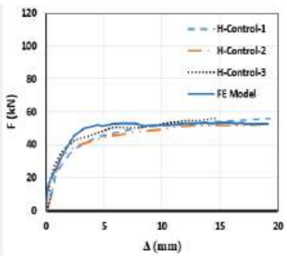
5. MODEL VALIDATION

5. 1. Control Beams The load–midspan deflection curves for control beams from experiments and FE analysis are plotted in Figure 4. The behavior of beams predicted by FE analysis is slightly stiffer and stronger after yielding point due to the assumption of fully bonded between concrete and reinforcement. However, the numerical results for control beams showed a very good agreement with the experimental results in term of load deflection curves. This agreement indicates that the constitutive models used for concrete and reinforcement were effectively captured the flexural behavior of RC beams. The concrete cracks that exhibited in the experimental and numerical models were compatible. The failure mode for both control beams was indicated by yielding of the tension reinforcement followed by concrete crushing at the mid-span section as experimentally observed in Figures 5a and 5b.

Its worthy to mention that the finite element model shows the plastic strains in the tension face of the beam in a separate view from the compression face. The compression face damage is shown Figure 5c. Where

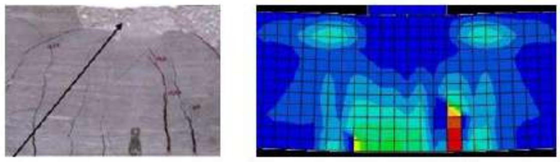


(a)

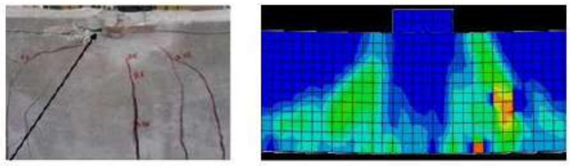


(b)

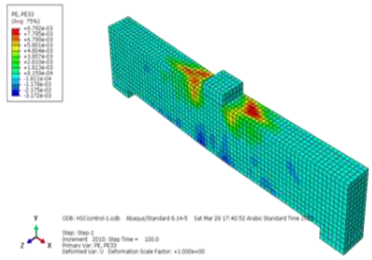
Figure 4. Comparison between experimental results and FE results for control beams: (a) L-control; and (b) H-control



(a)



(b)



(c)

Figure 5. Experimental/numerical failure modes for (a) L-control, (b) H-control, (C) compression face damage

the concrete crushed near the loading point in the compression fiber face.

The model was verified with additional work defined by Aljazaeri and Myers [2]. The finite element models were found to be compatible with the experimental test results. The ultimate load capacities for the modeled beams were compatible by 100% with the experimental tested control beam and 90% with the strengthened beam. However, the ultimate displacements for the modeled beams were compatible by 100% with the experimental tested control beam and 80% with the strengthened beam.

5. 2. Strengthened Beams The cohesive bond model was selected to simulate the load-midspan deflection responses of strenfthened beam since the previous studies proved that the perfect bond model was not able to simulate that accurately. The load–midspan deflection curves for strengthened beams from experiments and FE presented in Figure 6. All modeled beams showed very close results of yielding and ultimate loads in comparison with the experimental ones. For the purpose of relability, the FE model was also verifed based on Aljazaeri and Myers work [2] and close results were determined.

A comparison between the numerical and experimental results at the yielding stage and ultimate stage in terms of loads and deflections are given in Tables 2 and 3.

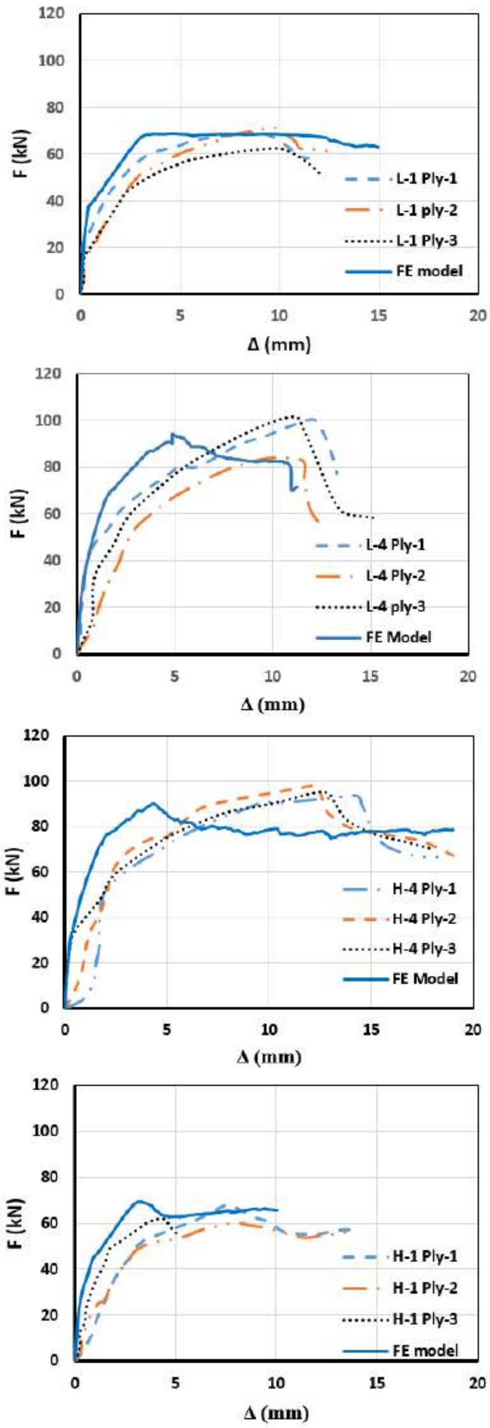


Figure. 6. Comparison between experimental results* and FE results of modified CBMs for strengthened beams: (a) L-1ply; (b) L-4ply; (c) H-1ply; and (d) H-4ply

The CBM precisely simulated the experimental load-midspan deflection response of both low and high strength concrete. For the beam with low compressive strength and strengthened with one layer (L-1ply), the ultimate load, initial stiffness and ultimate deflection were in excellent agreement with the corresponding experimental results. As well, the modeled beam (L-1ply) failed by slippage of the PBO-fiber as observed experimentally. The strengthened beam with 4 layers (L-4ply) (Figure 6b), the modeled beam was underestimated the ultimate load by 15%. However, the descending curve represented the debonding failure in the FRCM plies as exhibited experimentally.

TABLE 2. Test results: experimental and numerical load

Specimen ID	Experimental		Numerical		Difference	
	P_y (kN)	P_u kN)	P_y (kN)	P_u (kN)	$P_{y,(num.)}$ / $P_{y,(exp.)}$	$P_{u,(num.)}$ / $P_{u,(exp.)}$
L-control	38	51.4	39	54	1.03	1.05
L-1ply	57	67.7	57	68.9	1.00	1.02
L-4ply	75	99	77	84.5	1.03	0.85
H-control	40	55.8	42	53.5	1.05	0.96
H-1ply	54	63	58	69.6	1.07	1.10
H-4ply	78	96.8	81	90	1.04	0.93

TABLE 3. Experimental and numerical midspan deflection

Specimen ID	Experimental*		Numerical		Difference	
	δ_y (mm)	δ_u (mm)	δ_y (mm)	δ_u (mm)	$\delta_{y,(Num.)}$ $/\delta_{y,(Exp.)}$	$\delta_{u,(Num.)}$ $/\delta_{u,(Exp.)}$
L-control	1.4	25	1.9	25	1.36	1.00
L-1ply	2.3	12	2	12	0.87	1.00
L-4ply	3.3	12.2	3.3	4.4	1.00	0.36
H-control	1.6	24	2	24	1.25	1.00
H-1ply	2.4	5.5	2	3.5	0.83	0.64
H-4ply	3.8	15	2.8	4.5	0.74	0.30

The comparison between experimental and FE load-midspan deflection curves for high strength concrete strengthened with one and four layers of FRCM are presented in Figures 6c and 6d. It can be seen that the FE model was able to represent the strengthened beam behavior with a good level of accuracy. The ultimate load values predicted from FE models were within 15% and 7 % as a maximum divergence from the experimental values for beams strengthen with one and four layers, respectively. However, stiffer load-midspan deflection responses tolerated the yielding and ultimate deflections specifically for strengthened beams with four plies.

6. FAILURE MODES AND STRAIN MEASUREMENTS

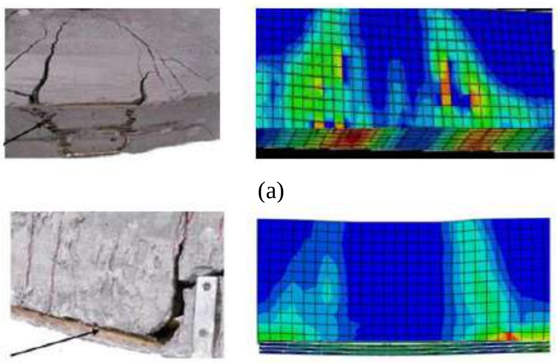
In order to increase the reliability of the FE model, the local measurement of a strain gage at the mid-span of beam, and observing mode of failures were compared with the corresponding FE value. The CBM was able to present the experimental failure modes of control and strengthened beams. The control beams detected a yielding of steel reinforcement with ductile failure as experimentally observed [3]. The strengthened beam with one FRCM ply failed by yielding of tension steel reinforcement followed by PBO-fibers' slippage (Figure 7a). While debonding of FRCM plies were detected in strengthened beams with four plies (Figure 7b).

The strain measurements of tested beams perceived by CBM were compared with the experimental results. Table 4 presented the numerical strain measurements for concrete, internal steel reinforcement, and FRCM composite at the beams' midspan.

All the modeled beams observed strain measurement of concrete that was equal approximately to the experimental ultimate strain (0.003 mm/mm). The numerical strain measurements of the internal steel reinforcement for the control and strengthened beams were agreed with the experimental ones. The numerical strain of the internal reinforcement ranged between (0.006 mm/mm to 0.009 mm/mm) for strengthened beams and (0.025 mm/mm to 0.07 mm/mm) for control beams, as observed experimentally [3].

However, the numerical strain measurements of FRCM composite were not compatible with the experimental ones. Since the experimental strain measurement was referred to the maximum strain in the PBO-fiber, but the numerical strain measurement referred to the maximum strain in FRCM composite as a laminate. In summary, it is observed that the CBMs revealed close numerical results to the experimental results for strengthened beams with one ply. In contrast, the numerical results of strengthened beams with four plies were far from the experimental results. The reason behind that the CBMs were dependent on tested specimens with one FRCM ply so CBMs did not account for four FRCM plies.

The numerical results found here were agreed with the previous studies' results reported by Elghazy et al.



(a)

(b)

Figure. 7. Experimental/numerical failure modes: (a) L-1ply; and (b) L-4ply

TABLE 4. Numerical strain measurements

Specimen ID	Strain at midspan, mm/mm		
	Concrete	Steel	FRCM composite
L-control	0.003	0.07	-
L-1ply	0.002	0.008	0.001
L-4ply	0.001	0.006	0.002
H-control	0.003	0.025	-
H-1ply	0.005	0.008	0.001
H-4ply	0.004	0.009	0.003

[12] where the ultimate loads and failure mechanism can be captured by CBMs but still, the numerical load-midspan deflection responses were varied from the experimental ones. Thus, new bond-slip laws need to address the increase in FRCM plies and denote precisely the cohesive damage parameters in the normal (mode I) and shear (mode II, III) directions.

7. PARAMETRIC STUDY

A comprehensive parametric study was conducted using the validated model. A total of 10 new models was used to study the effects of key variables that expected to have an essential impact on the behavior of strengthened RC beams. These parameters were the longitudinal steel reinforcement ratio and concrete compressive strength. For each key variable, both control and strengthened models were constructed in order to evaluate the contribution of open mesh fiber on the examined parameters. The observations for different parameters are described as follow:

7. 1. Longitudinal Steel Reinforcement Ratio A total of four FE models were conducted to evaluate the effect of steel reinforcement ratio (ρ) on the flexural behavior of RC beams strengthened with one layer of PBO fiber. For all models, concrete with normal compressive strength was used and four different sizes of steel reinforcement bars (2#12, 2#16, 2#20 and 2#25 mm bars) that equivalent to steel reinforcement ratio (ρ) of 0.24, 0.43, 0.67, and 1.05%; respectively.

The developed FE models results are presented in Table 5. In addition, Figure 8 compares the improvement of ultimate load of the strengthened beams with the control beams for different steel reinforcement ratio. As the reinforcement ratio increased a significant enhancement from strengthening can be achieved.

The results showed that all specimens failed by slippage in PBO also the post crack stiffness was increased with the increasing of steel reinforcement ratio which led to increase ductility of the specimens.

TABLE 5. Effect of different steel reinforcement ratio

Steel reinforcement ratio (%)	Ultimate load, Pu (kN)	Ultimate deflection (mm)	Post crack stiffness, k (kN/mm)
0.24	89	6.34	15.4
0.43	114	8.65	17.3
0.67	123	5.25	24
1.05	126	5.62	39

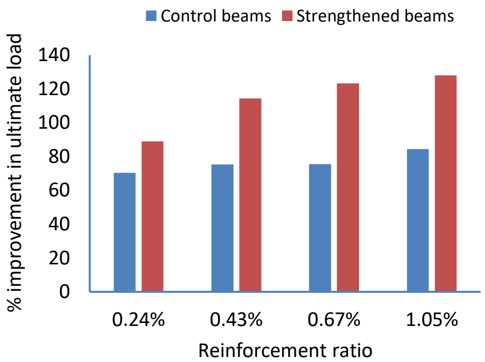


Figure 8. The improvement of ultimate load in different steel reinforcement ratio.

The FE results for flexural strengthened beams are compared with the analytical approach considering the ACI 549.4R code [9] in order to evaluate this approach under different steel reinforcement ratios as shown in Table 6. Based on ACI 549.4R [9], the flexural strength is calculated in accordance with Equation (4)

$$M_n = M_s + M_f \quad (4)$$

where M_n is the nominal flexural capacity of strengthened beam, and M_s and M_f are the contribution of internal steel reinforcement and external advanced composite material.

From Table 6, its observed that the results from the analytical model of ACI 549.4R [9] was very close to the results of FE models. The difference between two approaches within the range 12-25%. The reason behind this difference is the analytical model supposed uncoupled contribution of the steel and fiber which is not the case in FE models.

7. 2. Concrete Compressive Strength A total of six FE models were developed to study the effect of concrete compressive strength on the flexural behavior of strengthened beam. All beams had a constant external open mesh fiber (one layer of PBO) and internal reinforcement ratio of 0.24%. The concrete compressive strength was varied from 20 MPa to 50 MPa. The FE results are summarized in Table 7.

TABLE 6. Comparison between FE and analytical results

Steel reinforcement ratio (%)	ultimate load (kN) based on FE	ultimate load (kN) based on analytical method
0.24	89.05	78.81
0.43	114.33	112.09
0.67	123.35	170.34

TABLE 7. Effect of different concrete compressive strength

Concrete compressive strength (MPa)	Ultimate load, Pu (kN)	% Pu increase over control beams	Ultimate deflection (mm)	Uncracked stiffness (kN/mm)
20	66.5		3.8	85
25	67.7	102%	4	87
30	68.9	104%	5	89
35	70	105%	4.7	91
40	70	105%	3.1	94
50	71.4	107%	2.9	100

The load-deflection response was similar for simulated beams with concrete having compressive strength between 20 MPa to 35 MPa. Then, for simulated beams with higher compressive strength, load-deflection response become very brittle. A lower stiffness could be noticed for beams with low concrete strengths. However, the percentage increase in the ultimate loads was limited between 4% and 7%. The reason behind that is due to the fact that the flexural capacity of concrete beams reinforced with under reinforcement amount is dominated by the steel bar yield strength rather than compressive strength of concrete.

The concrete strength influenced the mode of failure in the FRCM strengthened system. For low concrete strength, a slippage of the PBO fiber was observed in the strengthened beams. While strengthened beams having high concrete strength above (35 MPa) observed a debonding failure mode in the FRCM system.

8. CONCLUSIONS

The FE models were developed in this study to analyze RC beams strengthened with FRCM composite under flexural load. The first step of modeling was to consider the cohesive bond models. Based on the result, the cohesive bond model presented good results in term of bond-slip behavior.

After validation with six full-scale RC beam tests, featuring different types of concrete, and different number of open mesh layers, in addition to different 10 other models that was used in a parametric study such as: different longitudinal steel reinforcement ratio and concrete compressive strength. The results showed a good agreement between the developed models and experimental results from uncracked stage till failure. The following conclusions can be drawn here:

1. The developed FE models were able to reasonably predict the flexural performance of RC beams with and without external strengthening with open mesh fibers.
2. The cohesive bond models prove to represent the load-midspan deflection responses of strengthened beams in comparison with their peers that were obtained from experimental work. The predicted yielding loads, ultimate loads, and failure modes are in excellent correlation with the experimental work.
3. The theoretical flexural capacity of strengthened beams based on ACI 549.4R code was evaluated based on numerical results. The developed models showed satisfactory computational capacity since the results of FE models and the theoretical results were close enough.
4. The variation in the concrete compressive strength has not dramatically influence the flexural behavior and the ultimate capacity of strengthened beam since failure mechanism was controlled by the steel bar yield strength rather than concrete compressive strength.

9. REFERENCES

1. Al-Jaberi, Zuhair A., John J. Myers, and Mohamed A. ElGawady. "Evaluation of FRP and FRCM composites for the strengthening of reinforced masonry walls." ***Journal of the American Concrete Institute,*** Vol. 327, (2018), 32-41. DOI: 10.14359/51702227.
2. Aljazaeri, Zena R., and John J. Myers. "Fatigue and flexural behavior of reinforced-concrete beams strengthened with fiber-reinforced cementitious matrix." ***Journal of Composites for Construction,*** Vol. 21, No. 1, (2017), DOI: 10.1061/(ASCE)CC.1943-5614.0000726.
3. Babaeidarabad, S., Loreto, G., Nanni, A. “Flexural strengthening of RC beams with an externally bonded fabric-reinforced cementitious matrix”, ***Journal of Composites for Construction***, Vol. 18, No. 5, (2014). DOI: 10.1061/(ASCE)CC.1943-5614.0000473.
4. D’Ambrisi, A., Focacci, F. Flexural strengthening of RC beams with cement-based composites. ***Journal of Composites for Construction,*** Vol. 15, No. 5, (2011) 707-720. DOI: 10.1061/(ASCE)CC.1943-5614.0000218.
5. Al-Jaberi, Zuhair A., John J. Myers, and Mohamed A. ElGawady. “ Out-of-plane behavior of RM walls strengthed with FRCM composite or NSM with cementitious adhesive.” 9th International Conference on Fibre-Reinforced Polymer (FRP) Composites in Civil Engineering, CICE 2018, DOI: 10.14359/51702227.
6. Aljazaeri, Zena R., and John J. Myers. "Strengthening of reinforced-concrete beams in shear with a fabric-reinforced cementitious matrix." ***Journal of Composites for Construction,*** Vol. 21, No. 5, (2017), DOI: 10.1061/ (ASCE)CC.1943-5614.0000822.
7. Loreto, Giovanni, et al. "Performance of RC slab-type elements strengthened with fabric-reinforced cementitious-matrix composites." ***Journal of Composites for Construction,*** Vol. 18, No. 3, (2014), A4013003. DOI: 10.1061/(ASCE)CC.1943-5614.0000415.
8. Al-Jabari, Zuhair, John J. Myers, and Mohamed ElGawady. "Effectiveness of FRCM System in Strengthening Reinforced Masonry Walls Subjected to Cyclic Loading." IABSE Symposium Report. ***International Association for Bridge and Structural Engineering***, Vol. 109, No. 39, (2017), DOI: 10.1016/j.compstruct.2018.11.085.
9. ACI (American Concrete Institute). Guide to design and construction of externally bonded fabric-reinforced cementitious matrix (FRCM) systems for repair and strengthening concrete and masonry structures. ACI 549.4 R. (2013). Farmington Hills, MI: ACI.
10. Maghsoudi, A. A., Reza Rahgozar, and Seyed Hamid Hashemi. "Flexural testing of high strength reinforced concrete beams strengthened with CFRP sheets." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 22, No. 2 (2009), 131-146.
11. Ombres, Luciano, and Salvatore Verre. "Structural behaviour of fabric reinforced cementitious matrix (FRCM) strengthened concrete columns under eccentric loading." ***Composites Part B: Engineering,*** Vol. 75, (2015), 235-249, DOI:10.1016/j.compositesb.2015.01.042
12. Elghazy, M., El Refai, A., Ebead, U., Nanni, A. "Corrosion-damaged RC beams repaired with fabric-reinforced cementitious matrix." ***Journal of Composites for Construction,*** Vol. 22, No. 5, (2018), 04018039, DOI: 10.1061/(ASCE)CC.1943-5614.0000873.
13. Al-Jaberi, Zuhair, John Myers, and Mohamed ElGawady. "Flexural capacity of out-of-plane reinforced masonry walls strengthened with externally bonded (EB) FRP." 7th International Conference on Advanced Composite Materials in Bridges and Structures Vancouver, British Columbia, Canada. 2016, DOI: 10.1016/j.conbuildmat.2018.02.043.
14. Mohsenzadeh, Sajjad, Ahamd Maleki, and Mohammad Ali Yaghin. "Experimental and Numerical Study of Energy Absorption Capacity of Glass Reinforced SCC Beams." ***International Journal of Engineering***, ***Transactions C: Aspects***, Vol. 32, No. 12, (2019), 1733-1744, DOI: 10.5829/IJE.2019.32.12C.06.
15. Alferjani , M. B. S., Abdul Samada, A. A., Elrawaffb, B. S. , and Mohamad, N. "Experimental and theoretical investigation on shear strengthening of RC precraced continuous t-beams using CFRP strips." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 28, No. 5, (2015): 671-676, DOI: 10.5829/idosi.ije.2015.28.05b.04.
16. Al-Jaberi, Zuhair, John J. Myers, and Mohamed A. ElGawady. "Experimental and analytical approach for prediction of out-of-plane capacity of reinforced masonry walls strengthened with externally bonded FRP laminate." ***Journal of Composites for Construction,*** Vol. 23, No. 4, (2019): 04019026, DOI: 1061/(ASCE)CC.1943-5614.0000947.
17. Fadaee, Mohammad Javad, and Hamzeh Dehghani. "Reliabilty-based torsional design of reinforced concrete beams strengthened with CFRP laminate." ***International Journal of Engineering, TransactionsA: Basics,*** Vol. 26, No. 10, (2013), 1103-1110.
18. Zhang SS, and Teng JG. "Finite element analysis of end cover separation in RC beams strengthened in flexure with FRP." Engineering structures. Vol. 75, (2014), 550-560, DOI: 10.1016/j.engstruct.2014.06.031.
19. Ombres, Luciano. "Flexural analysis of reinforced concrete beams strengthened with a cement based high strength composite material." ***Composite Structures,*** Vol. 94, No. 1, (2011), 143-155, DOI: 10.1016/j.compstruct.2011.07.008.

20. D'Antino, T., Carloni, C., Sneed, L. H., Pellegrino, C. "Matrix–fiber bond behavior in PBO FRCM composites: A fracture mechanics approach." ***Engineering Fracture Mechanics,*** Vol. 117, (2014), 94-111, DOI:10.1016/j.engfracmech.2014.01.011.

21. Focacci, F., D'Antino, T., Carloni, C., Sneed, L. H., Pellegrino, C. "An indirect method to calibrate the interfacial cohesive material law for FRCM-concrete joints." ***Materials & Design,*** Vol. 128, (2017), 206-217, DOI:10.1016/j.matdes.2017.04.038.

22. Carloni, C., D'Antino, T., Sneed, L. H., Pellegrino, C. "Three-dimensional numerical modeling of single-lap direct shear tests of FRCM-concrete joints using a cohesive damaged contact approach." ***Journal of Composites for Construction,*** Vol. 22, No. 1, (2018), 04017048, DOI: 10.1061/(ASCE)CC.1943-5614.0000827.

23. D'Antino, T., Colombi, P., Carloni, C., Sneed, L. H. "Estimation of a matrix-fiber interface cohesive material law in FRCM-concrete jonts." ***Composite Structures,*** Vol. 193, (2018), 103-112, DOI:10.1016/j.compstruct.2018.03.005.

24. Zo, X., Sneed, L. H., D'Antino, T., Carloni, C. "Analytical Bond-SlipModel for Fiber-Reinforced Cementitious Matrix-Concrete Joints Based on Strain Measurements." ***Journal of Materials in Civil Engineering,*** Vol. 31, No. 11, (2019), 04019247, DOI: 10.1061/(ASCE)MT.1943-5533.0002855.

25. ACI (American Concrete Institute). Report on High-Strength Concrete. ACI 363 R. (2010). Farmington Hills, MI: ACI.

26. ABAQUS 6.14 [Computer software]. Dassault Systemes, Waltham, MA.

27. Genikomsou, Aikaterini S., and Maria Anna Polak. "Finite element analysis of punching shear of concrete slabs using damaged plasticity model in ABAQUS." ***Engineering Structures,*** Vol. 98, (2015), 38-48, DOI:10.1016/j.engstruct.2015.04.016.

28. Desayi, Prakash, and S. Krishnan. "Equation for the stress-strain curve of concrete." ***Journal Proceedings***, Vol. 61. No. 3. 1964.

Persian Abstract

چکیده

سازه های بتونی مقاوم در برابر ماتریس سیمانی تقویت شده با الیاف (FRCM) به دلیل عملکرد مکانیکی و دوام بسیار گسترده شده اند. با این حال ، رفتار اعضای بتن آرمه تقویت شده با FRC تحت بارهای سرویس هنوز نگران کننده است و باید تلاش های بیشتری انجام شود. در این مطالعه ، یک مدل غیر خطی المان محدود سه بعدی (FE) برای مطالعه عملکرد تیرهای RC تقویت شده توسط FRCM ساخته شده است. مدل در برابر نتایج تجربی جمع آوری شده از شش پرتو آزمایش شده تحت خمش سه نقطه ای تأیید شد. در نتیجه ، پارامترهای اولیه عددی مورد مطالعه نسبت تقویت طولی فولاد و مقاومت فشاری بتن بودند. پارامترهای آسیب منسجم برای نشان دادن نتایج تجربی مورد بررسی قرار گرفت. همچنین ، ظرفیت خمشی نظری تیرهای تقویت شده بر اساس کد ACI-549.4R بر اساس نتایج روش عددی ارزیابی شد. به عنوان یک نتیجه گیری ، نتایج عددی در مورد بازده بار ، بار نهایی و حالت خرابی با نتایج تجربی مطابقت بسیار خوبی دارند. علاوه بر این ، مدلهای توسعه یافته از مطالعات پارامتریک تأثیر ناچیز مقاومت فشاری بتن بر افزایش ظرفیت نهایی تیر تقویت شده را نتیجه گرفتند. با این حال ، نسبت تقویت فولاد تأثیر عمده ای در افزایش ظرفیت نهایی تیرهای تقویت شده دارد.

IJE TRANSACTIONS A:Basics Vol. 34, No. 04, (April 2021) 803-810

International Journal of Engineering

Journal Homepage: www.ije.ir

Improved Strength and Stiffness Characteristics of Cross-laminated Poplar Timber Columns

J. A. Bhat*

Department of Civil Engineering, National Institute of Technology Srinagar, Hazratbal-190006, Jammu and Kashmir, India

PAPER INFO

Paper history:
Received 29 August 2020
Received in revised form 15 November 2020
Accepted 10 December 2020

Keywords:
Cold-formed Steel
Columns
Cross-laminated Timber
Fibre Reinforced Polymer
Poplar
Strength

ABSTRACT

Timber apart from being a cheap construction material possesses numerous environmental advantages, which makes it one of the highly sought construction material, particularly for moderately loaded residential structures. Timber due to the easy cultivation of timber trees can be made available in abundance. Thus, can serve as an efficient and sustainable building material, provided its structural potential is tapped fully. In this study, various performance improvement techniques have been used for enhancing the axial strength characteristics of a timber specie (Poplar), that is available in abundance, in the northern part of India. Different binding/wrapping techniques have been adopted to utilize Poplar in a laminated form, known as cross-laminated timber (CLT). It has been found that a strength improvement of about 20% can be achieved in CLT short columns by simply bolting the laminates together, while as this improvement can be as high as 32%, provided cold form steel (CFS) sheets are used for strengthening these CLT timber columns. Similarly, in the case of long CLT columns, a strength improvement of 50% is attained when a double cross helix of fibre reinforced polymer (FRP) cloth is adopted for the strengthening purpose. Furthermore, this study is aimed at utilizing small unused, otherwise, waste timber logs/pieces in columns with strength improvement techniques for improved axial strength performance.

***doi**: 10.5829/ije.2021.34.04a.06*

1. INTRODUCTION

The last 60 years have witnessed substantial growth in the strengthening of primary structural timber components like beams and columns. Initially, the strengthening was carried out primarily by using metallic reinforcement, including steel bars, pre-stressed stranded cables and bonded steel and aluminum plates [1]. Later, the metallic reinforcement was replaced by fibre-reinforced polymers (FRP) due to its high strength and stiffness features that would suit a variety of structural strengthening requirements. Many researchers have reported about the improvement in structural parameters after reinforcing the timber beams and columns with fibres and fibre-reinforced polymers (FRP) [2–7]. All these studies have indicated an improvement in the strength as well as stiffness characteristics in glulam timber elements. Further, the results revealed that the improvement in the stiffness as well as the strength can range between 40-60%.

Cross-Laminated Timber (CLT) an engineered wood, was first developed in Germany and Austria in the 1970s and 1980s, [8]. However, it was only after the mid-1990s, extensive research was carried out on the same and it was established that CLT has a variety of utility in the building construction [9–13]. Further, CLT has found its adaptability in timber–concrete composites, [14–16] and connections in timber structures as well, [17–20]. Even in seismically prone areas, CLT is an excellent material for mid and high-rise building construction [21, 22]. Also, CLT has displayed a satisfactory seismic-resistant performance, when used as floor diaphragms and shear walls [23, 24]. Recently, CLT has been used for tall timber and hybrid buildings, and seismic retrofitting of capacity deficit member [25]. Most of the research on CLT (glulam) has been focused on beams, while very limited research output on columns is available. To enhance the strength parameters, glue-laminated (glulam) timber columns were strengthened with FRP sheets [26–29]. Similarly, the

*Corresponding Author Institutional Email: *bhat_javed@nitsri.ac.in* (J. A. Bhat)

Please cite this article as: J. A. Bhat, Improved Strength and Stiffness Characteristics of Cross-laminated Poplar Timber Columns, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 803-810

feasibility of glue-laminated timber columns strengthened with FRP sheets for improved buckling resistance of columns was carried out by Taheri et al. [30], wherein, many parameters such as slenderness ratio, boundary conditions, FRP reinforcement length, and relative cost were investigated. A study was also carried out to improve the compression resistance of timber columns (with longitudinal cracks) by retrofitting (with FRP sheets), [27]. It was found that FRP laminates could offer an incremental increase in both the strength and stiffness.

Since locally available timber species can be utilized sustainably, benefitting regions all over the world, research was focused on locally available timber used conventionally in beams/columns, reinforced with fibres and FRP materials. Plevris and Triantafillou [2], Triantafillou and Deskovic [3] explored the role of carbon/epoxy fibre reinforced plastics in reinforcing fire wood products. The effect of pre-stressed carbon/epoxy FRP (CFRP) reinforcement bonded to European Beech Lumber was studied by Triantafillou and Deskovic [3]. Some studies were also carried out on Yellow-Poplar glulam beams reinforced on the tension side with glass/vinylester FRP [31] and with pultruded glass/vinylester FRP on the tension side and both on tension and compression sides. Similarly, Tingley and Leichti [32] discussed glulam made from lower grade Ponderosa Pine, reinforced in the tension zone with pultruded kevlar and carbon FRP.

A lower grade timber (Poplar) is locally available in almost all parts of Kashmir valley (a part of northern India). The timber from Poplar belongs to the specie, Aspen, and has currently very less structural importance, which is mainly used in roof trusses of low-cost houses, formworks of slabs and beams, apple boxes, ply boards, etc. Despite, Poplar timber being available in abundance in the above stated region, it has failed to achieve the deserved popularity to be used in a structural system and is so due to its low strength, moderate stiffness, and moderate-high shrinkage issues. An effort has been made to improve the strength of this type of timber so that its utilization as a constructional material can boost low-cost sustainable construction in the region. This study focuses on the utilization of Poplar timber in cross-laminated form, with various strength improvement techniques. To ensure the preparation of high-quality CLT, the availability of suitable raw material (local/domestic wood species) is essential. Further, it is also important to consider the strength properties, bonding properties, and durability of that particular species. Typically, temperate softwood lumber is used to manufacture CLT [33]. This makes Poplar wood suitable for CLT production and an effort has been made to study the axial strength behavior of externally strengthened CLT columns of locally available Poplar timber, with various strength improvement techniques.

As already stated the timber from popular has very less structural importance and is principally used in making apple boxes, plywood boards and roof trusses of low-cost houses in this region. Since cross laminating timber has the advantage that it randomizes the weak areas of timber and knots of timber throughout the volume of the column, hence increases the strength of weak timbers. Also, small unused timber pieces, otherwise waste-timber, can be utilized in structural elements (CLT columns) with strength improvement techniques. The novelty of this study is that with an increase in strength and stiffness, the utility of local Poplar timber will get enhanced at a low cost. Further, with an increase in the popularity of the use of timber in structural elements, industrial units can be established for mass production, which in turn will increase the employment and boost the economy in this industry deficit state.

2. MATERIAL AND MATERIAL DATA

The desired physical and mechanical properties of Poplar timber have been computed as per the procedure laid down in IS 1708: 1986, [34] by testing timber prisms of size 20mm x 20mm x 100mm. The material properties are given in Table 1. A typical sample of a sawn Poplar timber is shown in Figure 1.

In addition to the main material (timber), the other materials used for strengthening purposes are bolts, cold form steel sheets, and Carbon Fibre Reinforced Plastic (CFRP) strips. A brief description of these materials is given as follows:

TABLE 1. Physical and mechanical properties of Poplar timber

No.	Parameter	Average test value
1	Moisture content	17.00 %
2	Compressive strength parallel to the grain	27.0 N/mm^2
3	Compressive strength perpendicular to the grain	1.10 N/mm^2

Figure 1. A typical sample of sawn Poplar timber

2. 1. Bolts In CLT there is a tendency that the individual laminations of timber may rip apart. To hold the CLT laminations, three bolts of M6.8 (6mm dia) have been used. The properties of these bolts as provided by the supplier/manufacturer are shown in Table 2. Adequate spacing of 2d (200mm) between the bolts and an edge distance of 50mm has been provided. This would ensure the structural integrity of the CLT specimens and would prevent their delamination under the influence of the axial loading.

2. 2. Cold-formed Steel The columns were strengthened by cold-formed steel plates of 14G thickness (1.6mm) wrapped around all the long faces. Cold-formed steel conforming to Indian Standard IS-2062 [35] were used in this study. The properties of the cold-formed steel are given in Table 3.

2. 3. Carbon Fibre Reinforced Polymer (CFRP) The unidirectional CFRP strips of thickness 5mm were used to strengthen the timber CLT columns. These FRP strips were aimed to confine the timber laminations and improve the buckling/shear resistance of the CLT columns. As reported by the manufacturer, the modulus of elasticity, the tensile strength, and the strain at breakage for FRP are 175 kN/mm^2, 2000 N/mm^2 and 1.65%, respectively.

2. 4. Epoxy Epoxy was used as an adhesive between the timber laminations. The epoxy was a two-component adhesive one being resin and another hardener, where both the components were mixed in a ratio of 1:1. After thoroughly mixing the two, the mixed product was applied on the Poplar timber pieces within three minutes as prescribed in the manufacturer's manual provided by the manufacturer. As per this manufacturer's manual the epoxy was to be left to dry and harden for at least 24 hours before any testing could be done.

TABLE 2. Properties of Bolts

Designation	M6.8
Nominal Tensile Strength	600 N/mm^2
Minimum Brinell hardness HB	181
Nominal Yield Strength	480 N/mm^2
Elongation after fracture	8%

TABLE 3. Properties of cold steel

Material designation	CS sheets
Thickness	14G (1.6 mm)
Nominal Yield strength	250N/mm^2
Nominal Ultimate Strength	410 N/mm^2
Nominal %age elongation	23%

3. TEST SPECIMENS AND TESTING PROCEDURE

The sawn timber was kept for seasoning in air until the moisture content was reduced below 20%. The solid column was sawn into three laminations and the laminations were glued together with epoxy (araldite). The laminations were glued such that the grains in adjacent laminations are oriented perpendicular with respect to each other.

For the strengthening of columns, five different techniques were used. For each type of strengthening technique, six specimens were fabricated (three short columns and three slender columns), making a total of 30 CLT columns; of which 15 were short columns and 15 were slender columns. In addition, 06 control columns without any strengthening and laminates (single piece) were tested to lay the basis for the performance comparison of the various strengthening techniques adopted. The columns were tested under concentric axial loading conditions in the structural Laboratory at the National Institute of Technology Srinagar.

The section and lengths of the columns used are as follows:

a. Short Columns
Length of column = 500 mm;
Cross-section of the column = 100 mm x 100 mm;
l/d ratio = 5;
Cross-section of one laminate = 33 mm x 100 mm, three layers of laminate making an approximate cross-section 100 mm x 100 mm.

b. Long Columns
Length of column = 1000 mm;
Cross-section of the column = 75 mm x 75 mm;
l/d ratio = 13.33;
Cross-section of one laminate = 25 mm x 75 mm, three layers of laminate making approximate cross-section of 75 mm x 75 mm.

As stated above a total of thirty CLT columns and six single piece control columns of two categories were fabricated under similar environmental conditions. The description of each column and their designation is given in Table 4. The symbol ‘S’ and ‘L’ associated with CLT indicates a short column and long column respectively. A sample of these columns is shown in Figure 2.

Each column was tested to failure under concentric axial loading. Short columns with both ends pin were tested in a computerized universal testing machine (UTM) of 1000kN capacity (Figure 3), the loads and corresponding displacements are recorded in the computer through data control and acquisition system. Due to larger lengths, long columns could not fit in the UTM, and as such were tested in a loading frame of 1000 kN capacity with similar conditions. The load was applied through a hydraulic jack and the displacements were recorded with

TABLE 4. Description and Designation of columns

No.	Description	Designation	
		Short column	Long column
1.	Control column with no laminates	SCLT-0	LCLT-0
2.	CLT column joined with epoxy	SCLT-E	LCLT-E
3.	CLT column strengthened with bolts	SCLT-B	LCLT-B
4.	CLT column strengthened with cold steel.	SCLT-S	LCLT-S
5.	CLT column strengthened with single helix of FRP cloth	SCLT-SH	LCLT-SH
6.	CLT column strengthened with double cross helix of FRP cloth	SCLT-DH	LCLT-DH

(a) (b) (c) (d) (e) (f)

Figure 2. Sample Columns:(a) Solid column, (b) CLT Column, (c) Bolted CLT column, (d) Cold form sheet wrapped CLT column, (e) Single helix CFRP wrapped CLT column and (f) Double helix CFRP wrapped CLT column

the help of dial gauges fitted at the top and mid height of the column. The bottom and top surface of the columns were finished and leveled well to prevent any loading concentrations. Further, steel plates were used on top and bottom to distribute the load uniformly over the cross-section of the column.

4. RESULTS AND DISCUSSION

While performing the tests on columns, loads and vertical displacements were recorded during the entire testing period until their failure. These load and displacement values were averaged and plotted to obtain the stiffness, ultimate load, and maximum displacement. The failure pattern of short columns is shown in Figure 4. Control columns failed in the crushing of material, whereas

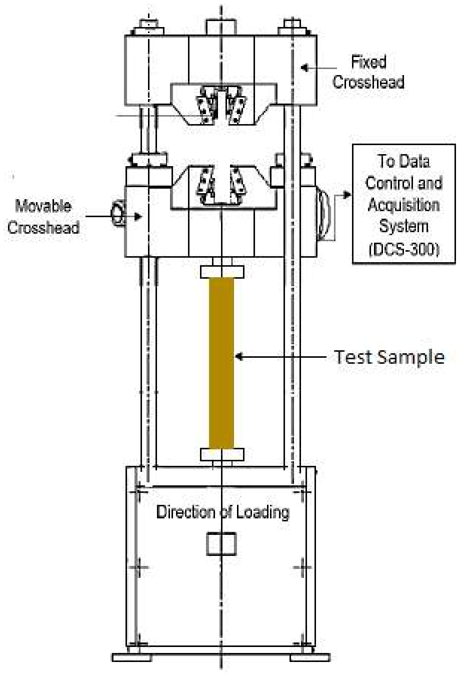


Figure 3. 1000 kN Universal Testing Macnine

SCLT-E failed due to the debonding of laminations Figures 4(a) and 4(b). Similarly, SCLT-B columns also failed due to the breaking of the Araldite bond followed by the buckling of the individual laminations Figure 4(c). However, SCLT-S columns failed by crushing of grains followed by the sudden opening of the cold form steel casing Figure 4(d). Similarly, in the case of SCLT-SH and SCLT-DH failure occurred due to the crushing of wood fibres and not by the debonding of the laminations as was found in the CLT column without strengthening (Figures 4(e) and 4(f)). Thus the FRP and cold steel sheet wrapping kept the laminations in position and enhances the axial carrying capacity of the CLT columns.

For short columns comparison of ultimate loads is shown in Figure 5. From this figure it is observed that the

(a) (b) (c) (d) (e) (f)

Figure 4. Failure pattern of short columns (a) SCLT-0, (b) SCLT-E, (c) SCLT-B, (d) SCLT-S, (e) SCLT-SH and (f) SCLT-DH

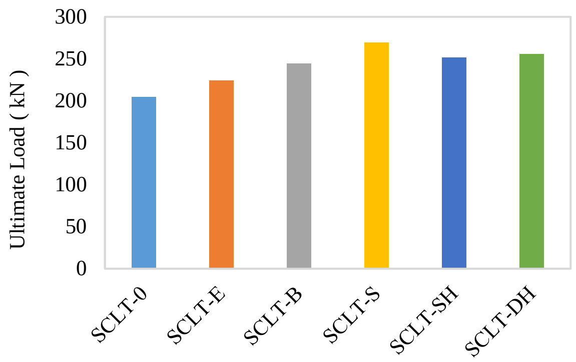


Figure 5. Comparison of Ultimate load carrying capacity for short columns

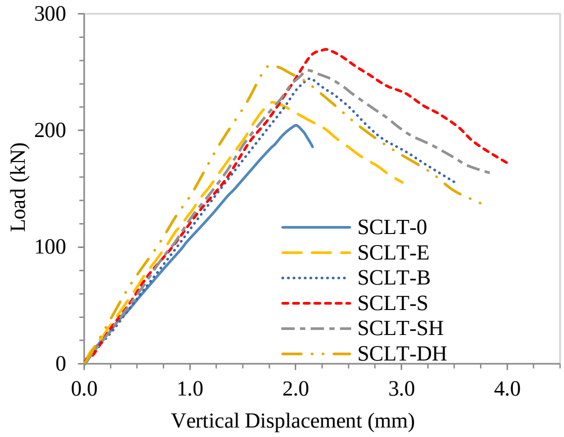


Figure 6. Comparison of load-displacement curves of various short columns

load-carrying capacity has increased in all the improvement techniques, however, SCLT-S has shown the maximum load-carrying capacity. The improvement in load-carrying capacity with different techniques is shown in Table 5. From this table, it is observed that using cross laminates with epoxy (SCLT-E) has increased the load-carrying capacity by 10%, and by adding few bolts (SCLT-B) the capacity has increased by about 20%. An increase in load-carrying capacity is as high as ~ 32% for SCLT-S i.e for the CLT column laminated with cold steel sheets. FRP helix has improved the load-carrying capacity by about 23% for SCLT-SH and 25% for SCLT-DH. Although there is a slightly lesser increase in strength (20%) in the case of SCLT-B; however, this is the simplest and cheapest technique for such kind of CLT columns.

Further, load-displacement curves for various SCLT columns have been plotted and a comparison of these load-displacement curves is shown in Figure 6. From this figure, it is observed that both stiffness and ductility of these CLT columns increases as compared to SCLT-0. The maximum stiffness and ductility are observed in the case of SCLT-DH due to confinement by FRP.

Similarly, long columns were tested, loads and corresponding vertical displacements were recorded till failure and a typical failure mode of these columns is shown in Figure 7. From this figure, it is observed that the control column failed by elastic buckling followed by

TABLE 5. Increase in load-carrying capacity in short columns

No.	Designation	Ultimate load (kN)	Compressive stress (N/mm^2)	% Increase
1	SCLT-0	204.24	20.42	0
2	SCLT-E	224.12	22.41	9.73
3	SCLT-B	244.10	24.41	19.52
4	SCLT-S	269.10	26.91	31.76
5	SCLT-SH	251.50	25.15	23.14
6	SCLT-DH	255.64	25.56	25.17

(a) (b) (c) (d) (e)

Figure 7. Failure patterns of long columns (a) LCLT-0, (b) LCLT-B, (c) LCLT-S, (d) LCLT-SH and (e) LCLT-DH

the cracking of timber on the tension side of the specimen, Figure 7(a). LCLT-E columns failed immediately due to the debonding of laminates and fell apart resulted in a brittle failure as such no image has been captured. Similar to control columns, LCLT-B columns failed due to the combined buckling and crushing of timber fibres, Figure 7(b). The specimen initially buckled under the axial load followed by the crushing of fibres on the compression side in the middle area. In the case of LCLT-S columns, the failure occurred due to initial elastic buckling followed by the debonding of steel casing from the timber surface and finally cracking of timber fibres on the compression side, Figure 7(c). Similarly, in the case of LCLT-SH and LCLT-DH, the columns initially buckled elastically followed by the crushing of timber fibres on the compression side in the portion that was not covered by FRP (Figures 7(d) and 7(e)). Due to the wrapping arrangement of FRP cloth around the cross-laminated column, increased the capacity of the columns with large deformations; thus, resulted in a delayed failure. As expected, contrast to the short columns all the long columns failed in buckling first followed by crushing of timber fibres.

A comparison of ultimate loads is shown in Figure 8. From this figure it is observed that for the column with cross laminates with epoxy only failed due to separation of laminates due to small buckling effects because of the slenderness effect. Further, it is observed that the compressive strength (overall load-carrying capacity) of all these columns is much lower than the short columns due to the slenderness effect as expected (Table 6). The improvement in load-carrying capacity with different techniques in the case of long columns is shown in Table 6. From this table, it is seen that the increase in load-carrying capacity is insignificant in case of LCLT-B (10%). However, an increase in load carrying capacity is as high as 50% in the case of the double helix CLT column (LCLT-DH). This increase in capacity in LCLT-DH owes to better confinement of laminates by cross FRP. Further, load-displacement curves for these long columns due to various techniques have been plotted and a comparison of these load-displacement curves is shown in Figure 9. From this figure, it is observed that stiffness and ductility increases only when there is sufficient confinement in the cross laminates of these columns, which is maximum in the case of LCLT-DH followed by LCLT-S. Further, it is also observed that LCLT-B has almost the same stiffness as that of LCLT-0 implying that effect of bolting in case of long columns is insignificant.

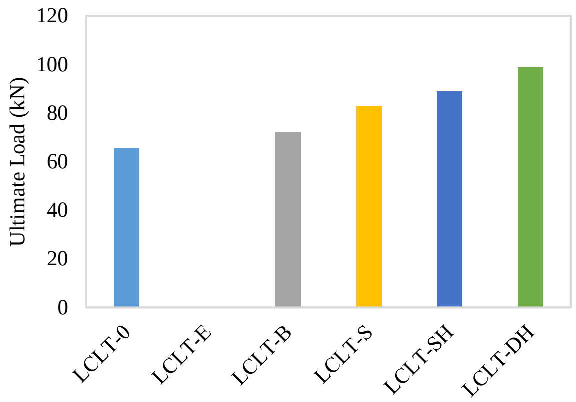


Figure 8. Comparison of Ultimate load carrying capacity for Long columns

TABLE 6. Increase in load-carrying capacity in long columns

No.	Designation	Ultimate load (kN)	Compressive stress (N/mm^2)	% Increase
1	LCLT-0	65.57	11.66	0
2	LCLT-E	---	0.00	----
3	LCLT-B	72.21	12.84	10.13
4	LCLT-S	83	14.76	26.58
5	LCLT-SH	88.81	15.79	35.44
6	LCLT-DH	98.77	17.56	50.63

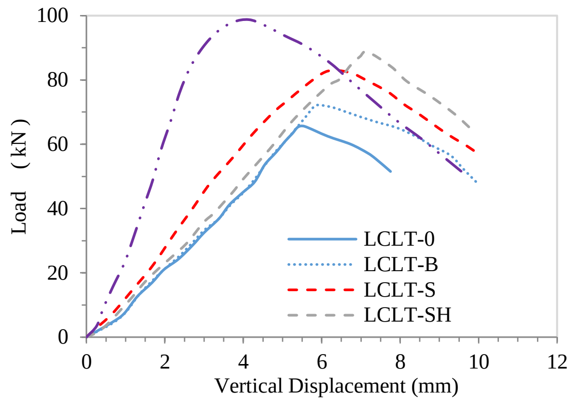


Figure 9. Comparison of load-displacement curves of various long columns

5. CONCLUSIONS

From the above discussion, it concluded that:

- The use of cross-laminated timber improves the strength, stiffness, and ductility in both short and long type columns.
- In the case of short columns, the improvement in load-carrying capacity and stiffness is more as compared to long columns.
- Simply using epoxy and few bolts in timber laminates increases (SCLT-B) the load-carrying capacity of columns by about 20%. However, in the case of cold form steel SCLT-S this increase is as high as 32%. Maximum stiffness and ductility are achieved in case of double cross helix FRP wrapping columns (SCLT-DH).
- In the case of long columns, the load-carrying capacity is derived by the slenderness of the columns, as such depends upon the confinement of these columns. In the present study maximum improvement in the load-carrying capacity, stiffness and ductility have been observed in the case of double cross helix FRP wrapping columns (LCLT-DH).
- The increase in load-carrying capacity is about 50% in these LCLT-DH columns.

6. ACKNOWLEDGMENT

The author acknowledges the Laboratory support provided by the National Institute of Technology Srinagar. Further, the author acknowledges the support provided by the Lab staff of the Structural Engineering Section of the Civil Engineering Department.

7. REFERENCES

1. Bulleit, W. M., Sandberg, L. B., and Woods, G. J., "Steel-Reinforced Glued Laminated Timber", ***Journal of Structural***

Engineering, Vol. 115, No. 2, (1989), 433–444. doi:10.1061/(ASCE)0733-9445(1989)115:2(433)

2. Plevris, N., and Triantafillou, T. C., "FRP-Reinforced Wood as Structural Material", ***Journal of Materials in Civil Engineering***, Vol. 4, No. 3, (1992), 300–317. doi:10.1061/(ASCE)0899-1561(1992)4:3(300)
3. Triantafillou, T. C., and Deskovic, N., "Prestressed FRP Sheets as External Reinforcement of Wood Members", ***Journal of Structural Engineering***, Vol. 118, No. 5, (1992), 1270–1284. doi:10.1061/(ASCE)0733-9445(1992)118:5(1270)
4. Yusof, A., and Saleh, A. L., "Flexural strengthening of timber beams using glass fibre reinforced polymer", ***Electronic Journal of Structural Engineering***, Vol. 10, (2010), 45–56.
5. Falk, R. H., and Colling, F., "Laminating Effects in Glued-Laminated Timber Beams", ***Journal of Structural Engineering***, Vol. 121, No. 12, (1995), 1857–1863. doi:10.1061/(ASCE)0733-9445(1995)121:12(1857)
6. Bhat, J., and Bhat, J., "Flexural Strengthening of Timber Beams using Carbon Fibre Reinforced Polymer Plates", ***International Journal of Civil Engineering and Technology***, Vol. 4, No. 5, (2013), 61–77.
7. Bhat, J., and Bhat, J., "Ductility of Timber Beams Strengthened using CFRP Pla", ***International Journal of Civil Engineering and Technology***, Vol. 4, No. 5, (2013), 42–54.
8. Schickhofer, G., and Guggenberger, W., "Elastic analysis of flexibly jointed laminated timber plates", 1st Congress of Croatian Society of Mechanics, (1994).
9. Fink, G., Kohler, J., and Brandner, R., "Application of European design principles to cross laminated timber", ***Engineering Structures***, Vol. 171, (2018), 934–943. doi:10.1016/j.engstruct.2018.02.081
10. Brandner, R., "Cross laminated timber (CLT) in compression perpendicular to plane: Testing, properties, design and recommendations for harmonizing design provisions for structural timber products", ***Engineering Structures***, Vol. 171, No. November 2017, (2018), 944–960. doi:10.1016/j.engstruct.2018.02.076
11. Lukacs, I., Björnfot, A., and Tomasi, R., "Strength and stiffness of cross-laminated timber (CLT) shear walls: State-of-the-art of analytical approaches", ***Engineering Structures***, Vol. 178, (2019), 136–147. doi:10.1016/j.engstruct.2018.05.126
12. Nolet, V., Casagrande, D., and Doudak, G., "Multipanel CLT shearwalls: an analytical methodology to predict the elastic-plastic behaviour", ***Engineering Structures***, Vol. 179, (2019), 640–654. doi:10.1016/j.engstruct.2018.11.017
13. Danielsson, H., and Serrano, E., "Cross laminated timber at in-plane beam loading – Prediction of shear stresses in crossing areas", ***Engineering Structures***, Vol. 171, (2018), 921–927. doi:10.1016/j.engstruct.2018.03.018
14. Schänzlin, J., and Fragiacomo, M., "Analytical derivation of the effective creep coefficients for timber-concrete composite structures", ***Engineering Structures***, Vol. 172, (2018), 432–439. doi:10.1016/j.engstruct.2018.05.056
15. Stepinac, M., Rajčić, V., and Barbalić, J., "Utjecaj dugotrajnog opterećenja na spregnute sustave drvo-beton", ***Journal of the Croatian Association of Civil Engineers***, Vol. 67, No. 3, (2015). doi:10.14256/JCE.1170.2014
16. Dias, A. M. P. G., Kuhlmann, U., Kudla, K., Mönch, S., and Dias, A. M. A., "Performance of dowel-type fasteners and notches for hybrid timber structures", ***Engineering Structures***, Vol. 171, (2018), 40–46. doi:10.1016/j.engstruct.2018.05.057
17. Jockwer, R., and Dietsch, P., "Review of design approaches and test results on brittle failure modes of connections loaded at an angle to the grain", ***Engineering Structures***, Vol. 171, (2018), 362–372. doi:10.1016/j.engstruct.2018.05.061
18. Sandhaas, C., and Görlacher, R., "Analysis of nail properties for joint design", ***Engineering Structures***, Vol. 173, (2018), 231–240. doi:10.1016/j.engstruct.2018.06.071
19. Cabrero, J. M., and Yurrita, M., "Performance assessment of existing models to predict brittle failure modes of steel-to-timber connections loaded parallel-to-grain with dowel-type fasteners", ***Engineering Structures***, Vol. 171, (2018), 895–910. doi:10.1016/j.engstruct.2018.03.037
20. Jockwer, R., Fink, G., and Köhler, J., "Assessment of the failure behaviour and reliability of timber connections with multiple dowel-type fasteners", ***Engineering Structures***, Vol. 172, (2018), 76–84. doi:10.1016/j.engstruct.2018.05.081
21. Popovski, M., and Gavric, I., "Performance of a 2-Story CLT House Subjected to Lateral Loads", ***Journal of Structural Engineering***, Vol. 142, No. 4, (2016). doi:10.1061/(ASCE)ST.1943-541X.0001315
22. Kuilen, J. W. G. V. De, Ceccotti, A., Xia, Z., and He, M., "Very Tall Wooden Buildings with Cross Laminated Timber", ***Procedia Engineering***, Vol. 14, (2011), 1621–1628. doi:10.1016/j.proeng.2011.07.204
23. Gavric, I., Fragiacomo, M., and Ceccotti, A., "Cyclic Behavior of CLT Wall Systems: Experimental Tests and Analytical Prediction Models", ***Journal of Structural Engineering***, Vol. 141, No. 11, (2015), 04015034. doi:10.1061/(ASCE)ST.1943-541X.0001246
24. Pei, S., van de Lindt, J. W., Popovski, M., Berman, J. W., Dolan, J. D., Ricles, J., Sause, R., Blomgren, H., and Rammer, D. R., "Cross-Laminated Timber for Seismic Regions: Progress and Challenges for Research and Implementation", ***Journal of Structural Engineering***, Vol. 142, No. 4, (2016). doi:10.1061/(ASCE)ST.1943-541X.0001192
25. Stepinac, M., Šušteršič, I., Gavrić, I., and Rajčić, V., "Seismic Design of Timber Buildings: Highlighted Challenges and Future Trends", ***Applied Sciences***, Vol. 10, No. 4, (2020), 1380. doi:10.3390/app10041380
26. Najm, H., Secaras, J., and Balaguru, P., "Compression Tests of Circular Timber Column Confined with Carbon Fibers Using Inorganic Matrix", ***Journal of Materials in Civil Engineering***, Vol. 19, No. 2, (2007), 198–204. doi:10.1061/(ASCE)0899-1561(2007)19:2(198)
27. Zhang, W., Song, X., Gu, X., and Tang, H., "Compressive Behavior of Longitudinally Cracked Timber Columns Retrofitted Using FRP Sheets", ***Journal of Structural Engineering***, Vol. 138, No. 1, (2012), 90–98. doi:10.1061/(ASCE)ST.1943-541X.0000423
28. Li, L., Yuan, S. L., Dong, J. F., and Wang, Q. Y., "An Experimental Study on the Axial Compressive Behavior of Timber Columns Strengthened by FRP Sheets with Different Wrapping Methods", ***Applied Mechanics and Materials***, Vols 351–352, (2013), 1419–1422. doi:10.4028/www.scientific.net/AMM.351-352.1419
29. Li, H., Qiu, H., Zhao, Z., and Lu, Y., "Axial compression behaviour of retrofitted long timber columns", ***Advances in Structural Engineering***, Vol. 21, No. 3, (2018), 445–459. doi:10.1177/1369433217719984
30. Taheri, F., Nagaraj, M., and Khosravi, P., "Buckling response of glue-laminated columns reinforced with fiber-reinforced plastic sheets", ***Composite Structures***, Vol. 88, No. 3, (2009), 481–490. doi:10.1016/j.compstruct.2008.05.013
31. Davlos, J, F., Salim, H, A., and Munipalle, U., "Glulam-GFRP Composite Beams for Stress-Laminated T-System Timber Bridges", Proceeding of the First International Conference on Advanced Composite Materials in Bridges and Structures. CSCE-CGC, Sherbrooke, Quebec, Canada, (1992), 455–463.
32. D.A. Tingley, R. J. L., "Reinforced Glulam: Improved Wood Utilization and Product Performance", Technical Forum-Globalization of Wood: Supply, Products, and Markets, Portland, or Forest Products Society, (1993).

33. Khairun, M., Uyup, A., Hamdan, H., Iskandar, M., and Anwar, U., "Cross-Laminated Timber: production of panel using sesenduk timber species", ***Timber Technology Bulletin***, Vol. 2016, No. 59, (2016), 1–6.
34. IS 1708: 1986 Reaffirmed 2005, "Technical Forum-Globalization of Wood: Supply, Products, and Markets, Portland, or Forest Products Society", Bureau of Indian Standards, India , 2005.
35. IS 2062: 2011, "Indian Standard for Hot Rolled Medium and High Tensile Structural Steel Specification (Seventh Revision)", Bureau of Indian Standards, India, 2011.

Persian Abstract

چکیده

الوار جدا از اینکه یک ماده ساختمانی ارزان قیمت است، دارای مزایای زیست محیطی بی‌شماری است که آن را به یکی از مصالح ساختمانی بسیار مورد پسند، به ویژه برای سازه‌های مسکونی با بار متوسط تبدیل می‌کند. الوار به دلیل کشت آسان درختان الوار می‌تواند به وفور در دسترس قرار گیرد، بنابراین می‌تواند به عنوان یک ماده ساختمانی کارآمد و پایدار عمل کند، به شرطی که از پتانسیل ساختاری آن کاملاً استفاده شود. در این مطالعه، از روش‌های مختلف بهبود عملکرد برای افزایش ویژگی‌های مقاومت محوری یک نوع چوب (صنوبر) استفاده شده است که به وفور در قسمت شمالی هند موجود است. تکنیک‌های مختلف اتصال / بسته‌بندی برای استفاده از صنوبر به شکل چند لایه، معروف به چوب چند لایه (CLT) استفاده شده است. مشخص شده است که می‌توان در ستون‌های کوتاه CLT به سادگی با پیچ و مهره ورقه ورقه‌ها، به بهبود مقاومت در حدود ۲۰ درصد دست یافت، در حالی که چون این بهبود می‌تواند تا ۳۲ درصد نیز برسد، به شرطی که از ورق‌های فولاد به فرم سرد (CFS) برای تقویت این ستون‌های چوبی استفاده شود. به طور مشابه، در مورد ستون‌های بلند CLT، هنگامی که یک مارپیچ دو طرفه از پارچه پلیمر تقویت شده با الیاف (FRP) برای هدف تقویت استفاده می‌شود، بهبود قدرت ۵۰ درصد حاصل می‌شود. علاوه بر این، این مطالعه با هدف استفاده از برگه‌های کوچک بلا استفاده و در غیر این صورت، ضایعات / قطعه‌های چوبی در ستون‌ها با تکنیک‌های بهبود قدرت برای بهبود عملکرد مقاومت محوری استفاده می‌شود.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 811-824

International Journal of Engineering

Journal Homepage: www.ije.ir

A Hybrid Modified Grasshopper Optimization Algorithm and Genetic Algorithm to Detect and Prevent DDoS Attacks

S. Mohammadi*, M. Babagoli

Department of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran

PAPER INFO

Paper history:
Received 14 October 2020
Received in revised form 23 December 2020
Accepted 28 February 2021

Keywords:
DDoS Detection
Cyber-security
Grasshopper Optimization Algorithm
Random Forest

A B S T R A C T

Cyber security has turned into a brutal and vicious environment due to the expansion of cyber-threats and cyberbullying. Distributed Denial of Service (DDoS) is a network menace that compromises victims' resources promptly. Considering the significant role of optimization algorithms in the highly accurate and adaptive detection of network attacks, the present study has proposed Hybrid Modified Grasshopper Optimization algorithm and Genetic Algorithm (HMGOGA) to detect and prevent DDoS attacks. HMGOGA overcomes conventional GOA drawbacks like low convergence speed and getting stuck in local optimum. In this paper, the proposed algorithm is used to detect DDoS attacks through the combined nonlinear regression (NR)-sigmoid model simulation. In order to serve this purpose, initially, the most important features in the network packages are extracted using the Random Forest (RF) method. By removing 55 irrelevant features out of a total of 77, the selected ones play a key role in the proposed model's performance. To affirm the efficiency, the high correlation of the selected features was measured with Decision Tree (DT). Subsequently, the HMGOGA is trained with benchmark cost functions and another proposed cost function that enabling it to detect malicious traffic properly. The usability of the proposed model is evaluated by comparing with two benchmark functions (Sphere and Ackley function). The experimental results have proved that HMGOGA based on NR-sigmoid outperforms other implemented models and conventional GOA methods with 99.90% and 99.34% train and test accuracy, respectively

doi: *10.5829/ije.2021.34.04a.07*

1. INTRODUCTION

The last decades have witnessed a revolution of Internet usage in many domains like e-commerce, e-government and so on. The expansion of the Internet is accompanied by the intensification of security violation issues. Denial of Service attack (DoS) is an intimidating attack which targets servers, online resources and network bandwidth. Victim's resources such as processors, bandwidth, database, memory, etc. are occupied with packet flooding which is generated by a malicious person or bot [1]. The devastation of servers or causing interruptions in online services is considered as the principal purpose of this attack. Distributed denial of services attack (DDoS) emerged as a powerful version of DoS with the capability of inflicting more destructive damage in a shorter span of time. Typically, DoS attacks are launched using one computer and one internet connection, whereas DDoS attacks are carried out by using several compromised computers (bots) and internet connections. Figure 1 shows one type of DDoS attack with multiple bots. In this figure, masters and slaves are hired in conjunction with an attacker to generate an enormous amount of packet [2, 3].

1. 1. DDoS Classification In DDoS attacks, the malicious user hires a network of zombie computers to incapacitate a server or website. DDoS attacks are categorized into three main groups: volume based attacks, protocol attacks and application layer attacks. Volume based attacks is the most common type of the aforementioned groups. These attacks send a large amount of requests or data to the victim's server with

*Corresponding Author Institutional Email: *Mohammadi@kntu.ac.ir* (S. Mohammadi)

Please cite this article as: S Mohammadi, M Babagoli, A Hybrid Modified Grasshopper Optimization Algorithm and Genetic Algorithm to Detect and Prevent DDoS Attacks, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 811-824

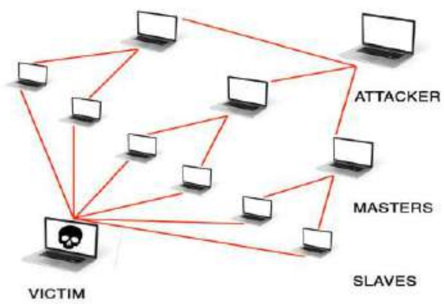


Figure 1. DDoS attack procedure

the purpose of overwhelming its bandwidth capability. Unavailability is considered as a major consequence of this type of attacks. Volume based attacks are prevalent in light of their simple amplifications; then, script kiddies can utilize this method for harming specific web services. Unlike the volume based attacks, Protocol features are abused in Protocol attacks [4]. What is employed in this type of DDoS attacks is an attempt to destruct or suspend a web resource. Indeed, intermediate communication devices (like load balancers and firewalls) are targeted to disrupt the communication of a website and its server. On the other hand, zombies (bots) are utilized in application layer attacks (or a 7-layer attack) to penetrate a specific server using the application's vulnerabilities [5]. This type of attacks requires fewer resources in comparison with the mentioned types on the grounds that it focuses on specific application packets which are sent through normal HTTP requests. Consequently, detection of application layer attacks is considered to be a laborious procedure [6]. The classification of DDoS attacks is described in depth in Figure 2.

1. 2. DDoS Prevention and Detection

Despite there being a lot of DDoS detection and prevention methods, deterring such attacks effectively is far-fetched if not impossible. In fact, the mitigation of DDoS risk has been the main aim of researchers. On the other hand, tracing back to the source is impractical as a result of IP spoofing (IP address is forged), stateless nature of network and similarity to flash crowd [7]. Therefore, source attack identification in DDoS attacks is an onerous endeavor. The detection and prevention techniques are divided into 3 categories: trace back methods, entropy based detection and intrusion detection and prevention systems [8].

Trace back methods have enhanced routers and protocol capabilities to track packets and uncover the source of attack. This method is often costly and with low accuracy. Packet marketing scheme and IP trace back technique are two schemes of this method [9]. Entropy is a measure of information theory which scales randomness of packets on specific router in entropy based DDoS detection. Indeed, the changes of flow's (packets with same destination address) abnormality are measured using entropy and the alarm would be raised if the rate of entropy is large. Hence, by tracking the entropy variation, the source of package is obtained. Information distance is the next metric which is used for distinguishing DDoS attacks and flash crowd. Intrusion detection system (IDS) is used to monitor the web traffic and report any suspicious activity to the administrator and intrusion prevention system (IPS) is designed to detect and prevent the attacks together with analyzing the data flow [10]. The segmentation of DDoS detection methods is illustrated in Figure 3.

In order to prevent DDoS attacks, many researchers have proposed different methodologies which focus on detection, prevention and trace back. Nevertheless, the lack of considering the limitation of real-time problems, complexity and massive data is a critical issue in DDoS detection strategies. With the intention of solving the aforementioned problem, anomaly based detection methods are used to create a profile of the normal traffic and then, detect the unknown attacks. Machine learning techniques are used to model a reliable behavior in network domain as a reference, and then compare new

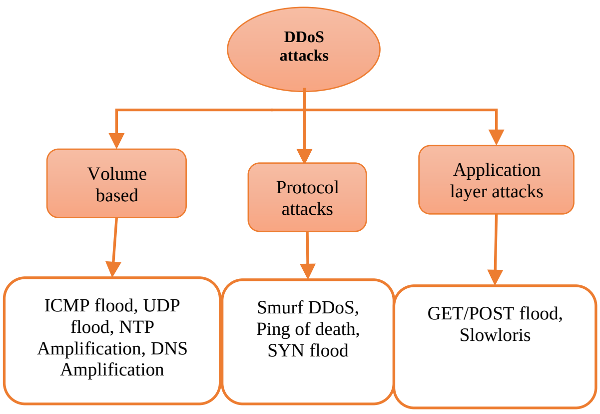


Figure 2. DDoS classification

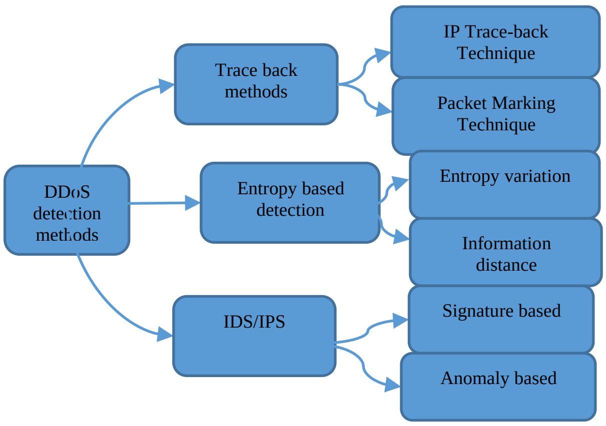


Figure 1. Classification of DDoS detection methods

ones with it. On the other hand, meta-heuristic algorithms are a nominated strategy to address the complexity and real-times issues and these algorithms can solve (NP)-hard problems [11]. additionally, some other important features like easy to use, cost-efficient and preparing important tools for both researchers and managers to solve the complex dilemma, makes these algorithms more popular [12]. Considering the no free lunch theory, there is no guarantee to one meta-heuristic algorithm outperforms in all problems. In order to reach better performance in a specific problem, there are several new meta-heuristic algorithms can be proposed or conventional algorithms can be modified or different algorithms can be combined with each other [13]. Accordingly, a combination of machine learning techniques and meta-heuristic algorithms can be used to boost the performance of detection method in terms of accuracy, speed and extendibility[14].

In this paper, the combination of machine learning and meta-heuristic algorithms is utilized to resolve the issue at hand and an efficient model is proposed to detect and prevent DDoS attacks. In order to evaluate the performance of the proposed method, two benchmark cost functions are applied into the model. An up-to-date dataset (CICIDS) is used to train and test the model. CICIDS consists of reliable and real-world samples which cover different attacks properly [15]. Ultimately, one machine learning technique: random forest (RF) and two meta-heuristic algorithms (Hybrid Modified Grasshopper Optimization algorithm and Genetic Algorithm (HMGOGA) and conventional Grasshopper optimization algorithm(GOA)) are utilized for feature selection and DDoS detection, respectively. The rest of this paper is organized as follows: section 2 is devoted to a review of related literature regarding the previous studies, section 3 touches upon the proposed DDoS detection method, section 4 describes and discusses the experimental results and finally, section 5 concludes the present research.

2. LITERATURE REVIEW

Taking into account the background of DDoS attacks, some major DDoS detection mechanisms are described in this section. The focus of this section is on machine learning and data mining approaches. Gu et al. [16] utilized the semi-supervised weighted k-mean and hybrid feature selection (SKM-HFS) method to detect DDoS attacks. In order to validate their experiments, they used three benchmark datasets and the results of their proposed mechanism were compared with one another. The feature selection performance was evaluated using TOPSIS method. As shown in their results, SKM-HFS had better performance in both time-consumption and precision. Finally, with the purpose of evaluating SKM-HFS in the real world, a real experimental environment was employed to appraise the functionality of the proposed algorithm. Like other experimented datasets, SKM-HFS has shown an acceptable performance in the real world dataset. Gharvirian et al.[6] used a perceptron neural network along with computing entropy of flow and flow initiation rate in order to detect the DDoS attack in the SDN controller. Indeed, In this research, the neural network makes improvement in the detection accuracy and false alarm rate and proves the existence of attack by investigating the 3 features of network traffic. Considering the vitality of the detection time, the proposed model used the neural network just for suspicious flows. The detection accuracy approximately reached 92% and the delay of detection obtained 23.55 seconds which is proof positive of the detection efficiency. Ghasemi et al.[17] proposed a multi-stage detection model and in each stage, they concentrated on one type of attack. They used genetic algorithm in order to select the most important features of each type of attack. In this paper, a novel chain model is proposed to detect each type of attack respectively. After one type of attack is detected, the chain model deletes specific labels from the dataset. In order to evaluate the proposed model, two benchmark datasets (NSL-KDD and KDD cup99) were used. The accuracy of average detections for two datasets were 97.5% and 98.9%, respectively. Four different classifiers are used as the fitness function for genetic algorithm, decision tree outperforms other methods in most cases. Nezhad et al. [18] have applied time series model and chaotic system to distinguish between legitimate and suspicious traffic. Two features (number of packet and number of source IP address) have been used as detection metric in every minute, and a detection accuracy of about 99% has been obtained. The Box-Cox transformation, Auto Regressive Integrated Moving Average (ARIMA) and Lyapunov were utilized for data processing, predicting and classification phases, respectively. Many DDoS detection methods based on machine learning were tried on SDN[2]s (Figure 4). Artificial neural network was employed to detect the different types of DDoS attacks [19]. ICMP flood, SYN flood, UDP flood and DNS reflection were experimented using proposed collaborating intrusion detection system (CIDS). The emulation results have proved the proficiency of ANN[3]-based CIDS in SDN. Conversely, some inherent features of SDN can be used to assist the confrontation with DDoS attacks. In this trend, SDNs advantages can be used for DDoS detection in cloud environment [20]. The methodology for DDoS defeating in SDN can use learning techniques (Machine learning/Deep learning)

[2] Software define network
[3] Artificial neural network

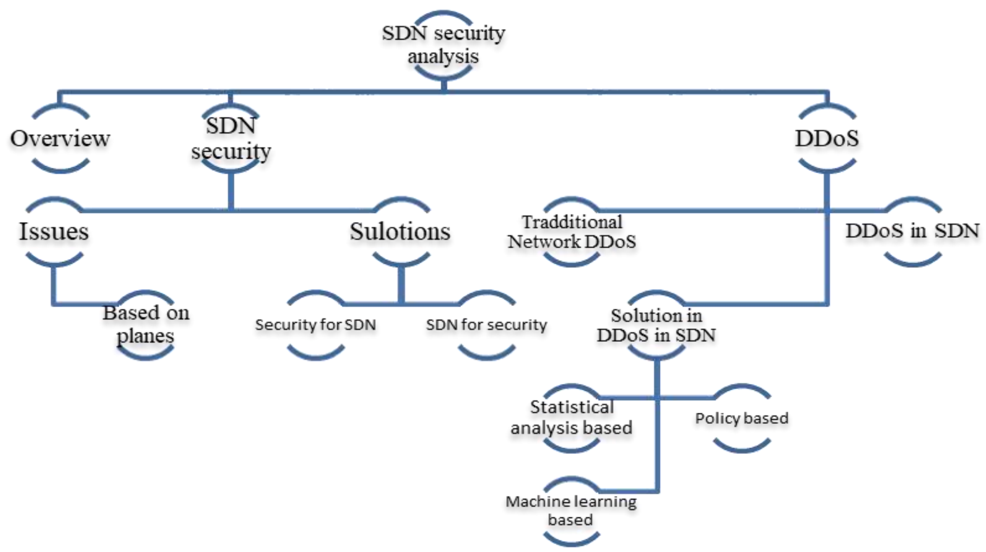


Figure 4. Research categories in SDN security domains [25]

to ameliorate detection rates and reduce the computation cost and time. Niyaz et al. [21] have proposed a network application on the basis of deep learning for multi-vector attack detection. Deep learning methods have been employed to remove irrelevant features and select the most important ones. Three implemented models were investigated for feature classification and the accuracy of 95.65% was obtained from SAE[4] (stacked sparse auto-encoders and soft-max classifier) approach. Arivudainambi et al. [22] have proposed Lion optimization algorithm [23], a new meta-heuristic algorithm, to detect DDoS attacks in SDN. The vector feature selection method has been applied to the selected dataset (NSL-KDD) to collect an appropriate feature subset and a combination of Lion Optimization Algorithm and Convolutional Neural Network has been used for training and testing. As it was demonstrated in their results, the average accuracy reached to 96%. Sreeram et al. [24] proposed a bio-inspired bat algorithm to detect HTTP flood attack in a short time frame and with high speed. The CAIDA dataset was used to select the most important feature for the proposed model. Afterwards, the selected features were used to train and test the bat algorithm. As shown in their results, they have obtained 94.8% accuracy in detecting HTTP flood attacks.

However, most of the researches mentioned above were incapable of adequately detecting new DDoS attacks at the right time. Some of the main drawbacks of the existing literature which were used as motivation for our research are as follows: lack of high accuracy accompanied by acceptable time-consumption and extendibility, difficulty in detection of unknown and zero-day DDoS attacks, lack of expansion of new methodology for detecting DDoS attacks, not using comprehensive datasets, etc.

[4] Stacked auto encoder

3. PROPOSED DDoS DETECTION MODEL

Having plenty of information in networks packet, abnormal behavior of packets can be recognized using analysis methods. Therefore, some available datasets are included in network data to provide efficient context for network security researches. NSL-KDD [26] and CICIDS [27] are the two most popular datasets that have been provided for network threats investigations. Due to antiquated data in NSL-KDD. This research has employed CICIDS in order to assess the proposed model. Some traffic features in CICDS are ineffectual, leading to degradation of learning quality, more memory consumption and an increase in computational time. Feature selection methods can properly solve these issues. In this paper, a machine learning method, RF, is used to collect more important features. Next, HMGOGA is utilized to detect DDoS attacks using the selected features. At last, a comparison of the conventional method and other research is made to evaluate the performance of our model.

3. 1. CICIDS Dataset The DDoS dataset applied in this manuscript is adopted from UNB repository [27]. The dataset consists of 77 features and one label column. The types of traffic are indicated using label column. Due to the problem of diversification in other datasets, CICIDS comprised 225,745 samples which include legitimate and attack traffic. The feature description is available by details in [15]. Table 1 provides the general information about CICIDS.

3. 2. Feature Selection Feature selection methods, a type of dimensional reduction techniques, are used to transform features into a new space with low dimensions. Indeed, the irrelevant features are eliminated from the set of features and the most important ones remain [28]. Prior to our DDoS detection method, RF is used to improve the detection throughput. RF as a popular machine learning method makes use of tree based decision making and results in an efficient performance regarding the low over-fitting, good predictive accuracy and ease of use [29]. The relevant features are selected by their impurity measures; as a matter of fact, when a tree is trained in RF, decrease of weighted impurity in a tree can be computed by each feature. Therefore, the average of each feature's impurity reduction can be used to rank the features in a forest. According to correlated features in CICIDS, the most important features have led to low impurity [30]. Continuing this process, selected features are used as feed for meta-heuristic algorithm and the procedure goes on to detect the DDoS attack. Figure 5 shows the framework of the proposed detection method.

3. 3. Proposed DDoS Detection Method In the detection phase, initially, the GOA is trained using the selected feature subset to develop an ability to detect unknown attacks. The GOA is a new meta-heuristic algorithm that is inspired by the behavior of grasshopper swarms while finding food and moving toward the source of food. The mathematical model of this algorithm is used for optimization problems [31]. The GOA is based upon swarm intelligence and population based categories. The merits of GOA algorithms are proved using several test functions and it is outperformed in cases of productivity from exploration to exploitation, randomness quality, search space coverage, scape from local minimum and fast convergence to optimum solution [32]. The mathematical model of GOA is described below. The position of each grasshopper is obtained using Equation (1).

$$X_i(t+1) = S_i(t) + G_i(t) + A_i(t) \tag{1}$$

where, S, G and A denote the Social Interaction (SI), gravity force and effect of wind flow, respectively.

The social interaction is the main parameter of GOA and plays a pivotal role in problem optimization. Social interaction is defined as follows:

$$S(i) = \sum_{j=1(j\neq i)}^{nPop} s(d_{ij})\hat{d}_{ij} \tag{2}$$

where, d_{ij} denotes the Euclidean distance between the i th and j -th grasshopper and $s(d_{ij})$ is a social force function that is based on attraction and repulsion forces. Hence, the effect of grasshoppers on each other is measured using this function.

$$s(d) = fe^{-\frac{d}{l}} - e^{-d} \tag{3}$$

where, f is gravity intensity and l is gravity length scale. By rewriting the main equation:

$$X_i = \sum_{j=1(j\neq i)}^{nPop} s(|x(i)-x(j)|)\frac{x(i)-x(j)}{d_{ij}} - ge_g - ue_w \tag{4}$$

where, e_g, e_w and u denote the unit vector across the direction to the center of the earth, unit vector across

TABLE 1. Details of CICIDS

	# of rows	# of column	# of legitimate traffic	# of attack traffic
CICIDS	225745	78	97718	128027

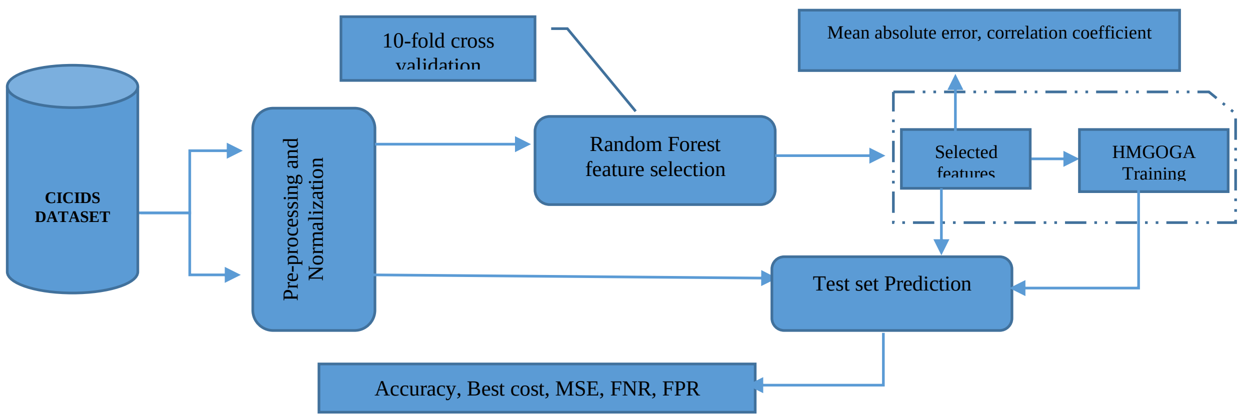


Figure 5. Proposed model framework

wind blow direction and fix drift, respectively. The proposed equation is not usable for optimization problems due to the weakness of exploration and exploitation in finding the optimal solution, thus utilizing the modified equation in this paper as follows:

$$X_i^d(t+1)=c\left\{\sum_{j=1(i\neq j)}^{nPop} c\frac{ub_d-lb_d}{2}s\left(\left|x_i(t)-x_j(t)\right|\right)\frac{x_i(t)-x_j(t)}{d_{ij}}\right\}+\hat{T}_d(t) \quad (5)$$

Where, ub and lb are the upper bound and lower bound in d-th dimension of Equation (6) and $\hat{T}(t)$ denotes the best solution that has been found so far.

GOA is useful for solving many complex global optimization problems. Nevertheless, there are some drawbacks in the conventional GOA like low convergence speed and stuck in the local minima [33]. Due to the complexity of our search space, the position of each grasshopper must update more accurately considering the whole search space. In order to reduce the time of finding the optimal solution and increase the convergence speed of GOA, a new SI strategy has been introduced in this paper. In the conventional GOA, the social interaction for each grasshopper can be obtained using the distance between one grasshopper and others. Indeed, in each iteration, the specific grasshopper can be affected by both far and close grasshoppers equally (Equations (2) and (3)). Consequently, improper effects of far grasshoppers cause an increase in computing time and algorithms iterations in order to find the optimal solution. In this paper, a novel strategy is introduced for SI which moderates grasshopper effects. Social interaction force is calculated for each grasshopper using just the nearest grasshoppers, not far ones. Indeed, by organizing the position of grasshoppers, the speed and power of finding global optimum is increased; however, sometimes this algorithms may gets stuck in local optima for complex optimization problem due to its weak diversity [34].

$$S(i)=\sum_{j=1(j\neq i)}^{nNearest} s(d_{ij})\hat{d}_{ij} \quad (6)$$

Unbalanced exploration and exploitation is another weakness of the original GOA that can lead into falling in a local optimum trap [35]. To overcome this obstacle, in this research, two genetic algorithm principles, crossover and mutation, are added to the GOA. Crossover and mutation operators in GA, work for diversification and intensification phases and one of the main characteristics of the GA algorithms is the behavior of operators that operates by chance. Although this characteristic is considered as negative point of GA, makes our model more powerful in the exploration phase. This proposed algorithm- called Hybrid Modified Grasshopper Optimization and Genetic Algorithm (HMGOGA)- is considered to be an extension of MGOA which enhances the exploration and exploitation power of the algorithm for the purpose of avoiding local minimums. In further detail, in each iteration, after the grasshopper position has been updated (Equation (5)), parents are selected from a new grasshopper population and offspring created by exchanging genes. Parent selection is randomly uses one of the following three methods in each iteration: Roulette Wheel, Random and tournament selection. Subsequently, binary crossover is applied to the selected parents and offspring can be created. Before adding the offspring to grasshopper population, in the mutation phase, some of the grasshopper genes are flipped randomly. The exploration and exploitation capabilities of modified GOA (MGOA) are improved using crossover and mutation, respectively. In order to fair operation of exploration and exploitation, the c parameter in Equation (5) is decreased by increasing iteration. The detailed Pseudo code for the proposed method can be described as follows and the flowchart of proposed algorithms is illustrated in Figure 6.

```
Start
Initialized parameters and population
for i=1:MaxIteration
        - update all grasshopper position (Eq.
        5). Social interaction for each
        grasshopper is calculated just by closer
        grasshoppers.
        - evaluate population using cost
        functions
        - generate random number between 1:3 to
        determine parents' selection strategy
        - apply crossover and create offspring
        - apply mutation for random grasshoppers'
        genes.
        - evaluate offspring using
        - concatenate created offspring and
        grasshopper population and select the
        best population.
End
Finish
```

In order to find near-optimal solution in meta-heuristic algorithms, parameter tuning is a major concern of researchers for improving efficiency and capability of algorithms. Parameter tuning provides more flexibility and robustness in problem solving and it requires careful initialization. Indeed, the parameter tunning is highly related to the complexity of the problems but many researchers propose an optimal value for key parameters of the algorithms [36]. In this research, after using trial and error method for finding best value in algorithm setting, the researcher's proposition is used. For instance, in order to fair usage of exploration and exploitation proportional in Equation (5), the c parameter is calculated as follows [35]:

$$c=c_{\max}-currentIt*\frac{c_{\max}-c_{\min}}{\max It} \quad (7)$$

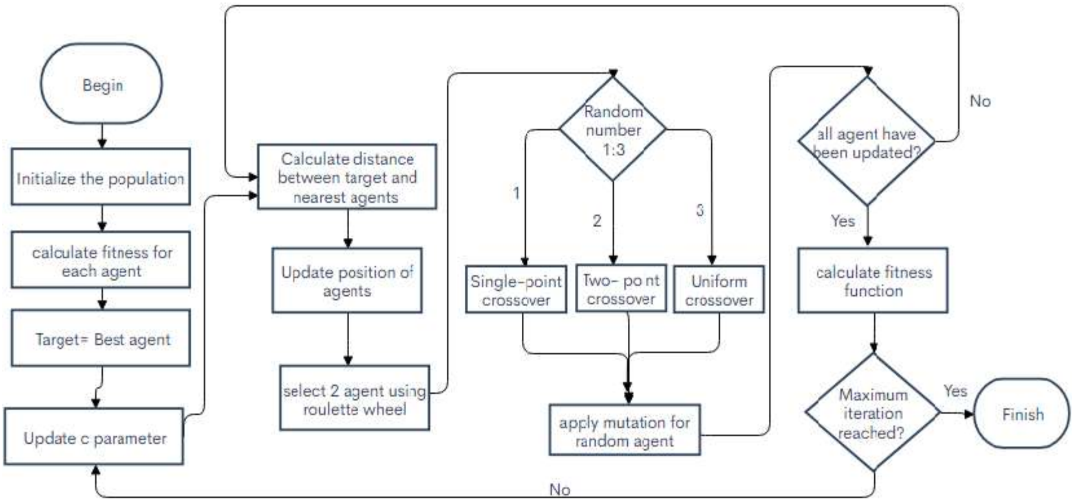


Figure 6. Flowchart of proposed model

where, $c_{\max}$ and $c_{\min}$ is maximum and minimum value, $currentIt$ is a current iteration and $\max It$ indicates the maximum iteration. According to Equation (7) the c parameters reduce in each iteration. In fact, the c parameter is updated to reduce exploration and increase exploitation ($c_{\max}$ and $c_{\min}$ are considered 1 and 1e-4, respectively).

In order to prove the efficiency of the proposed model, several benchmark functions are used as cost functions for GOA. In fact, the optimum coefficients of cost functions are calculated using the meta-heuristic algorithm. Three benchmark functions are applied so that the model performance can be figured out in various conditions such as Sphere, Ackley function [37, 38] and non-linear regression [39]. The sphere function is a simple continuous, convex and unimodal function which is widely used for optimization problems [40]. Ackley function is utilized as a more sophisticated function in the proposed model. Ackley function was first applied to genetic hill-climbing [38]. Ackley function is a non-convex function which is used for testing the optimization algorithms. Nonlinear regression is a form of linear regression analysis in which the relation between dependent and independent variables are nonlinear. Regression analysis mainly aims to model the observational data and find the relationship between responsible variables (y) and predictors (x). The relation between x and y is investigated using coefficients and the optimal values of the coefficients are obtained through using meta-heuristic algorithms. The implemented equations are shown in Table 2.

Where, α is the parameter that must be optimized, β is a random number between [-1, 1], x denotes the input vector which consists of the selected features, and n is the number of the selected features.

After the training phase, GOA predicts the label of test data by using cost functions and then, the accuracy of prediction is evaluated using Mean Square Error(MSE). The MSE formulation for both train and test phases is shown in Equation (8).

TABLE 2. Implemented benchmark cost function

Name	Cost functions	Equation number
Sphere	$y = \sum_{i=1}^{n} \alpha_i x_i$	(9)
Ackley function	$y = -20 \times exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} a_i x_i^{2}}) - exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi a_i x_i)) + 20 + exp(1)$	(10)
Combined Nonlinear regression-Sigmoid function	$y = \dfrac{1}{1+e^{-(\sum_{i=1}^{N}\alpha_i x_i + \sum_{j=1}^{N}\sum_{k=j+1}^{N}\alpha_{jk} x_j x_k + \beta)}}$	(11)

$$MSE = \frac{\sum_{z=1}^{N} (class\_label(z) - f(s))^2}{N} \qquad (8)$$

where, $f(s)$ shows the desire output, N represent the number of rows in the dataset and $class\_label(z)$ denotes the real class of each packet.

4. SIMULATION RESULT

In this section, HMGOGA algorithm is applied to several cost functions and the results have been thoroughly compared. In order to evaluate the proposed model, some credible research projects have been compared and the efficient application of our proposed method is investigated using DDoS detection. Firstly, data are cleaned and the null features are removed from CICIDS dataset in the pre-processing phase. Next, the data will be normalized for the purpose of homogenization of features effect (Equation (12)).

$$x' = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \qquad (12)$$

where, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of each feature, respectively. After normalization, 70% of data is randomly selected for training and the other 30% is kept and used for testing evaluation. In order to eliminate redundant features, improve detection accuracy and reduce the computational cost and required storage, RF is utilized as a feature selection technique. As it can be deducted from the results, 20 features are efficiently selected among 77 features and the performance of the selected ones is validated by Decision Tree (DT). Figure 7 shows the ranking of the selected features on the basis of their ranking merit and the details of these selected features are described in Table 3.

TABLE 1. Details of selected features

Number	Name	Merit
feature 44	ACK Flag Count	0.083998
feature 0	Destination Port	0.078377
feature 45	URG Flag Count	0.061353
feature 10	Bwd Packet Length Max	0.055670
feature 12	Bwd Packet Length Mean	0.054299
feature 50	Avg Bwd Segment Size	0.047392
feature 47	Down/Up Ratio	0.039466
feature 48	Average Packet Size	0.037646
feature 13	Bwd Packet Length Std	0.036543
feature 38	Packet Length Std	0.033712
feature 37	Packet Length Mean	0.033700
feature 36	Max Packet Length	0.027459
feature 8	Fwd Packet Length Mean	0.023853
feature 59	min_seg_size_forward	0.023748
feature 6	Fwd Packet Length Max	0.018635
feature 43	PSH Flag Count	0.017955
feature 49	Avg Fwd Segment Size	0.017868
feature 22	Fwd IAT Std	0.017700
feature 39	Packet Length Variance	0.017284

As shown in the results, the most valuable features are extracted efficiently and about 75% of the irrelevant features are removed from the subset. For assessing the feature subset using DT classifier, two measures are used: Mean Absolute Error (MAE) and Correlation Coefficient.

The distance between two variables is measured using MAE that is investigated in this phase to calculate the average absolute difference between the prediction and true class label values. The strength of relation between two variables is obtained by Correlation Coefficient metric. Indeed, the dependence of features to the labeled class is defined using correlation coefficient. The competency of feature subset is proved by high correlation and low MAE for which 96.84% and 3% were obtained, respectively.

Henceforth, the HMGOGA algorithm is qualified to detect DDoS attack using the most important features. In order to strike high performance strategy, different benchmark test functions are utilized and the parameters of functions are optimized to decrease the MSE and increase the detection accuracy. Primarily, like other meta-heuristic algorithms, a random population value between [-1, 1] is generated as coefficients of target functions. In each iteration the powerful particles (grasshoppers) are maintained and the weakest ones are eliminated. The strength of particles is defined using MSE. As a matter of fact, each row of population is multiplied into the target function using training dataset and the population is changed in each iteration according to HMGOGA procedure.

Finally, the most eligible particle is considered as an elected coefficient. In this step, 70% of data is selected randomly for training and the other 30% is considered as test data. Figure 8 (b, d, f) demonstrate the training phase of HMGOGA algorithm with different cost functions. As it can be observed in Figure 8, the downward trend of the MSE indicates the successful process of training. The performance of the proposed model is checked by predicting the precision of test data

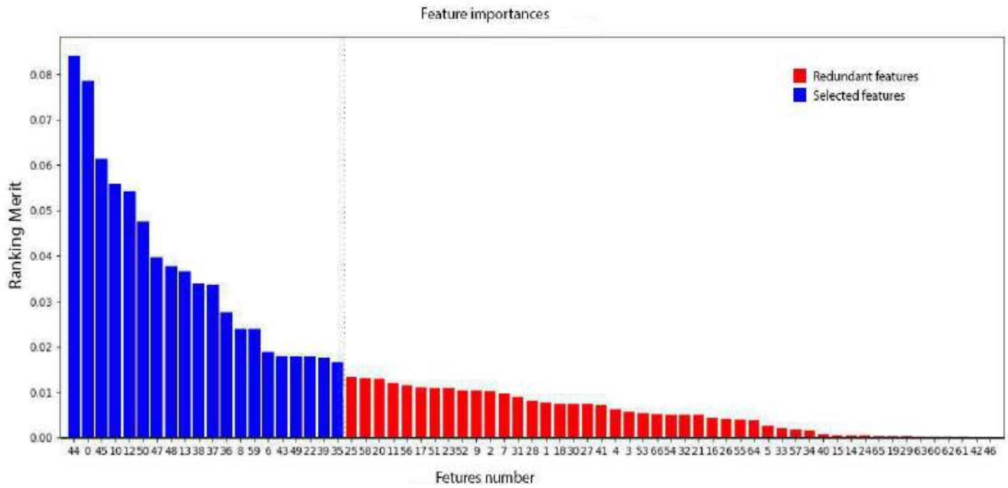


Figure 7. Features ranking

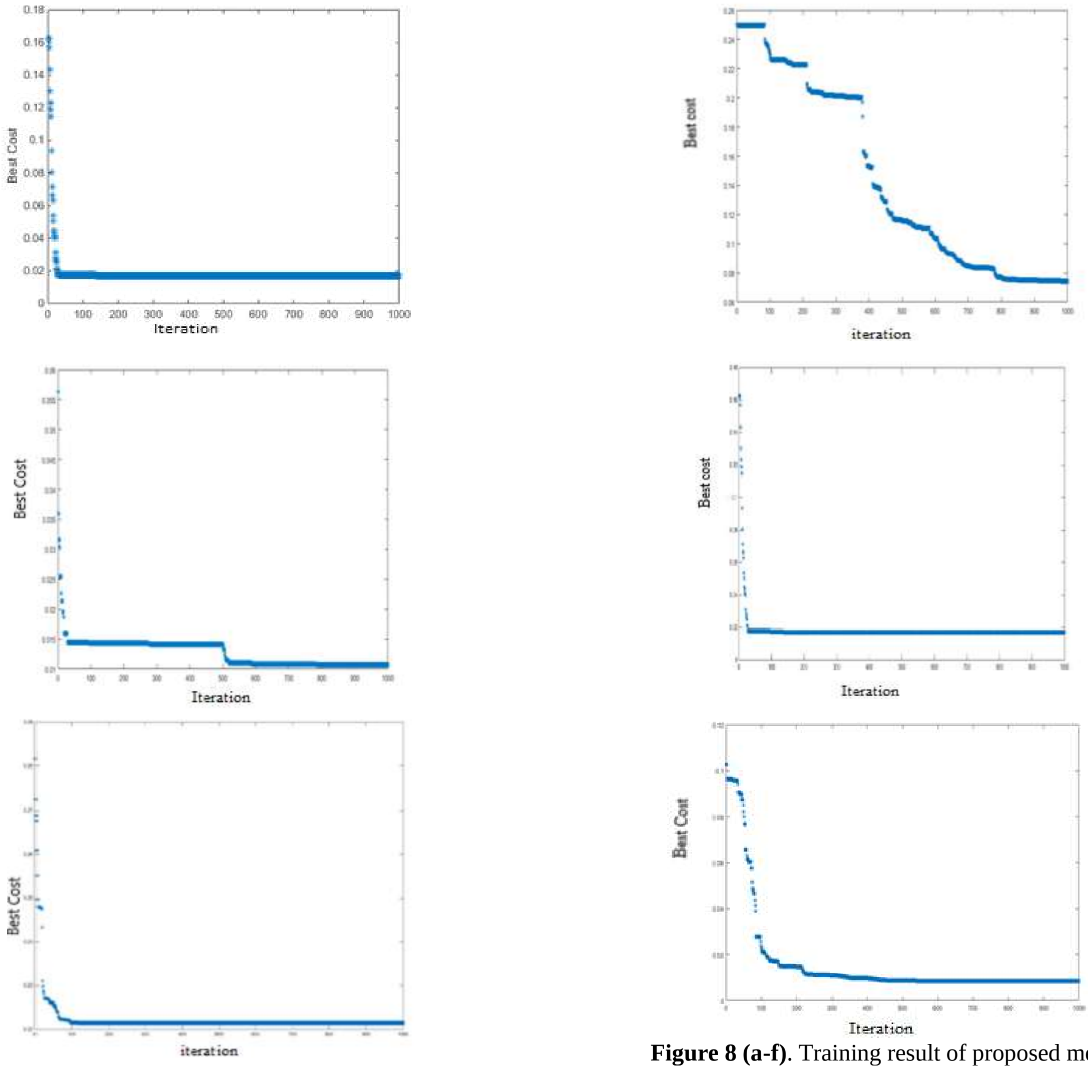


Figure 8 (a-f). Training result of proposed method label and confusion matrix (Equations (13)-(14)). Testing data includes unknown network packets which are classified as legitimate packets or DDoS attacks using trained HMGOGA, elected coefficients and considering cost function. At last, conventional GOA is implemented in similar conditions to make a direct comparison with the proposed model (Table 4).

$$Accuracy\,(train\,/\,test) = \frac{\sum_{i=1}^{N} N_{(predict_i = desire_i)}}{N_T} \times 100 \tag{13}$$

		Actual	
		Positive	**Negative**
Predicted	**Positive**	True Positive (TP)	False Positive (FP)
	Negative	False Negative(FN)	True Negative(TN)

Rate

$$TPR = \frac{TP}{TP+FN} \qquad FPR = \frac{FP}{FP+TN}$$
$$FNR = \frac{FN}{FN+TP} \qquad TNR = \frac{TN}{TN+FP} \tag{14}$$

According to detection sensitivity of DDoS attacks, the confusion matrix is used to prove the stability of the proposed method. The most important metrics in attack detection are TP and FN where TP is the number of attacks correctly classified as attacks and FN is the number of attacks incorrectly classified as normal records. Furthermore, TN and FP are the number of normal records correctly classified as normal records and number of normal records incorrectly classified as attacks (Equation (14)).

As shown in the results, HMGOGA with nonlinear regression cost function has converged efficiently and obtained high-performance accuracy with low FN. Therefore, the proposed model using non-linear cost function has a better performance in comparison with other cost functions. Additionally, HMGOGA outperforms conventional GOA algorithm in every aspect (Table 4). In order to depict the details of HMGOGA algorithm, Mean Cost, Best Cost and Worst Cost of all implemented models are obtained but due to space restriction in this paper, we have just illustrated one of them in Figure 9 to exhibit the different trends of the Worst, Mean and Best populations. According to this figure, the charts are not coincident with one another but all of the 3 charts have a downward trend after a specific iteration, for the generation of each population is based on the prior population.

As shown in Table 4, the proposed nonlinear regression fitted to the model better than other cost functions. The coefficients of cost function are optimized using the implemented meta-heuristic algorithms. The results suggested that HMGOGA based on proposed nonlinear regression has a more accurate and robust performance compared with conventional GOA in case of DDoS detection. The robustness of the proposed model is proved by obtaining low FP and FN.

The receiver operating characteristics (ROC) curve is one of the most important metrics for evaluating the model's performance and it can compare sensitivity versus specificity across a range of values for the ability to predict dichotomous outputs. The area under the ROC curve is another measure of test performance that is shown in Figure 10. The area under curve (AUC) in HMGOGA shows better performance compared to conventional GOA. Indeed, The AUC of HMGOGA depicts the high accuracy and high recall of the proposed model in different thresholds. In order to prove the robustness of the model, some other statistical

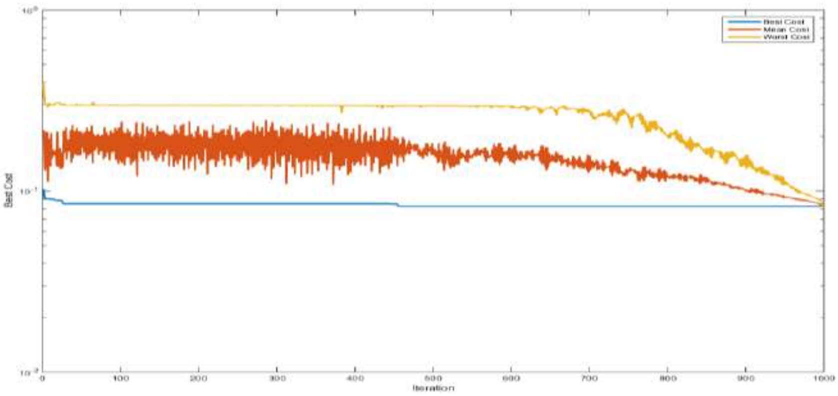


Figure 9. Best, Mean and worst cost of GOA for Ackley function

TABLE 2. Performance analysis

	Performance metrics	Sphere function	Ackley function	Non-linear regression
HMGOGA	Train accuracy:	98.930%	92.552%	99.907%
	Test accuracy:	97.375%	84.015%	**99.3496**%
	FN rate:	0.001	0.001	**0.001**
	FP rate:	0.042	0.264	**0.012**
	TP rate: (Sensivity or recall)	0.998	0.998	0.998
	TN rate:	0.957	0.735	0.988
GOA	Train accuracy:	97.459%	93.742%	98.277%
	Test accuracy:	92.773%	91.436%	95.846%
	FN rate:	0.001	0.038	0.001
	FP rate:	0.118	0.102	0.082
	TP rate: (Sensivity or recall)	0.998	0.961	0.998
	TN rate:	0.881	0.897	0.917

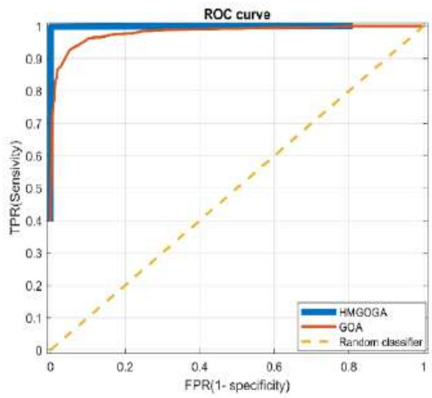


Figure 10. Roc curve for HMGOGA and GOA

test like confidence intervals are calculated in this paper. The robustness of the model is illustrated in Figure 9; where the best cost and mean cost are approximately converged to a single point [41]. According to Equation (15), the obtained accuracy is 99.35% 0.001 by 99% confidence interval (z=2.576).

$$ConfidenceInterval = z \times \sqrt{\frac{error \times (1 - error)}{N}} \quad (15)$$

Considering the related research projects on DDoS and intrusion detection systems, many researchers have employed machine learning and meta-heuristic techniques. Hence, some related studies are investigated to compare the performance of our proposed model and validate the efficiency. As shown in Table 5, our proposed method has utilized a novel dataset and meta-heuristic algorithm in DDoS detection scope and achieved a better detection accuracy in comparison with other related research.

TABLE 5. Comparison analysis

References	Dataset	Detection method	accuracy
Bista et al. [42]	CAIDA UCSD DRAPA 2000	Heuristic clustering algorithm and Nave-Bayesian classifier	99.45% , 86.73%
Arivudainambi et al. [22]	NSL-KDD cup	Lion optimization algorithm + Convolutional neural network	98.2%
Sreeram et al.[24]	CAIDA	bio-inspired bat algorithm	94.8%
Proposed method	**CICIDS**	**HMGOGA + Random Forest**	**99.3496%**

5. CONCLUSION

In this paper, a DDoS detection framework has been devised based on the latest meta-heuristic algorithm called GOA in conjunction with a new benchmark dataset called CICIDS and a potent feature selection method called Random Forest. Initially, the most relevant features are extracted from CICIDS dataset using RF feature selection method. The aforementioned dataset consists of 77 features about 75% of which are irrelevant features and are removed from the dataset. Selected features are utilized by GOA algorithm with different cost functions. Considering some weaknesses of GOA: low convergence speed and getting stuck in local minimum, this algorithm is modified and then combined with genetic algorithm (named HMGOGA). As it can be inferred from the results, HMGOGA algorithm confirms better performance in terms of accuracy and robustness. Regarding the novelty of the utilized dataset and meta-heuristic algorithm, the main contributions of this proposed framework is listed below.

1. The Random Forest (RF) feature selection method is applied to our utilized dataset and the 20 most important features among 77 are selected. The performance of the preferred feature subset is validated using DT classifier measures: Mean Absolute Error (MAE) and Correlation Coefficient. High correlation and low MAE have been obtained from our selected features.

2. Low convergence speed and getting stuck in local optimum are two drawbacks of GOA algorithm. In order to overcome these shortcomings, the new SI method is proposed to solve the convergence problem (MGOA). Then, Genetic algorithm is employed to adjust the exploration and exploitation phase and improve the search capability of GOA. The proposed algorithm is called HMGOGA.

3. Two meta-heuristic algorithms (HMGOGA and conventional GOA) are implemented to detect DDoS attacks. HMGOGA and GOA are implemented in similar conditions. The results indicate that the HMGOGA outperforms GOA in terms of detection accuracy and robustness.

4. In order to evaluate the performance and extendibility of the HMGOGA, the proposed framework is implemented using 3 benchmark functions: Sphere, Ackley function and the combined NR-Sigmoid function. The results reveal that NR-Sigmoid function proves to perform better in both HMGOGA and GOA by 99.34 and 95.84 percent test accuracy. In addition, the accuracy of HMGOGA is higher than GOA in all targeting functions. Indeed, nonlinear regression discovered the hidden relation of data more properly.

6. REFERENCES

1. A. Saied, R. E. Overill, and T. J. N. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks,", ***Neurocomputing*** Vol. 172, (2016), 385-393. https://doi.org/10.1016/j.neucom.2015.04.101
2. O. Osanaiye, K.-K. R. Choo, M. J. J. o. N. Dlodlo, and C. Applications, "Distributed denial of service (DDoS) resilience in cloud: Review and conceptual cloud DDoS mitigation framework," ***Journal of Network and Computer Applications*** Vol. 67, (2016), 147-165. https://doi.org/10.1016/j.jnca.2016.01.001
3. C. Kolias, G. Kambourakis, A. Stavrou, and J. J. C. Voas, "DDoS in the IoT: Mirai and other botnets," ***Computer***, Vol. 50, No. 7, (2017), 80-84. DOI: 10.1109/MC.2017.201
4. J. Wang, M. Yin, and J. J. T. C. S. Wu, "Two approximate algorithms for model counting," ***Theoretical Computer Science*** Vol. 657, (2017), 28-37. https://doi.org/10.1016/j.tcs.2016.04.047
5. C. Wang, T. T. Miu, X. Luo, J. J. I. T. o. I. F. Wang, and Security, "SkyShield: a sketch-based defense system against application layer DDoS attacks," ***IEEE Transactions on Information Forensics and Security,*** Vol. 13, No. 3, (2017), 559-573. DOI: 10.1109/TIFS.2017.2758754
6. Gharvirian, Fateme, and Ali Bohloli. "Neural network based protection of software defined network controller against distributed denial of service attacks." ***International Journal of Engineering***, ***Transactions B: Applications*** Vol. 30, No. 11 (2017), 1714-1722. DOI: 10.5829/ije.2017.30.11b.12
7. A. R. a. Yusof, N. I. Udzir, and A. J. I. J. o. D. E. T. Selamat, "Systematic literature review and taxonomy for DDoS attack detection and prediction,", ***International Journal of Digital Enterprise Technology***, Vol. 1, No. 3, (2019), 292-315. https://doi.org/10.1504/IJDET.2019.097849
8. Ö. Cepheli, S. Büyükçorak, G. J. J. o. E. Karabulut Kurt, and C. Engineering, "Hybrid intrusion detection system for ddos attacks," ***Journal of Electrical and Computer Engineering*** Vol. 2016, (2016). https://doi.org/10.1155/2016/1075648
9. M. H. Bhuyan, D. Bhattacharyya, J. K. J. S. Kalita, and C. Networks, "E-LDAT: a lightweight system for DDoS flooding attack detection and IP traceback using extended entropy metric," *Security and Communication Networks* Vol. 9, No. 16,(2016) 3251-3270. https://doi.org/10.1002/sec.1530
10. Shamshirband, Shahab, Mahdis Fathi, Anthony T. Chronopoulos, Antonio Montieri, Fabio Palumbo, and Antonio Pescapè. "Computational intelligence intrusion detection techniques in mobile cloud computing environments: Review, taxonomy, and open research issues." ***Journal of Information Security and Applications*** Vol. 55, (2020), 102582. https://doi.org/10.1016/j.jisa.2020.102582
11. Beiki, H., S. M. Seyedhosseini, V. R. Ghezavati, and S. M. Seyedaliakbar. "Multi-objective Optimization of Multi-vehicle Relief Logistics Considering Satisfaction Levels under Uncertainty." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 33, No. 5, (2020), 814-824. DOI: 10.5829/ije.2020.33.05b.13
12. Fathollahi-Fard, Amir Mohammad, Mostafa Hajiaghaei-Keshteli, and Reza Tavakkoli-Moghaddam. "Red deer algorithm (RDA): a new nature-inspired meta-heuristic." ***Soft Computing*** (2020), 1-29. https://doi.org/10.1007/s00500-020-04812-z
13. Hajiaghaei-Keshteli, Mostafa, Ahmad J Afshari, and Elahe Nasiri. "Addressing the freight consolidation and containerization problem by recent and hybridized meta-heuristic algorithms." ***International Journal of Engineering***, ***Transactions C: Aspects,*** Vol. 30, No. 3, (2017), 403-410. DOI: 10.5829/idosi.ije.2017.30.03c.10
14. Fathollahi-Fard, Amir Mohammad, Mostafa Hajiaghaei-Keshteli, and Seyedali Mirjalili. "A set of efficient heuristics for a home healthcare problem." ***Neural Computing and Applications,*** Vol. 32, No. 10, (2020), 6185-6205. https://doi.org/10.1007/s00521-019-04126-8
15. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "A Detailed Analysis of the CICIDS2017 Data Set," in *International Conference on Information Systems Security and Privacy*, 2018, 172-188: Springer. https://doi.org/10.1007/978-3-030-25109-3_9
16. Y. Gu, K. Li, Z. Guo, and Y. J. I. A. Wang, "Semi-Supervised K-Means DDoS Detection Method Using Hybrid Feature Selection Algorithm," ***IEEE Access*** Vol. 7, (2019), 64351-64365. DOI: 10.1109/ACCESS.2019.2917532
17. Ghasemi, J., and J. Esmaily. "A novel intrusion detection systems based on genetic algorithms-suggested features by the means of different permutations of labels' orders." ***International Journal of Engineering, Transactions A: Basics,*** Vol. 30, No. 10, (2017), 1494-1502. DOI: 10.5829/ije.2017.30.10a.10
18. S. M. T. Nezhad, M. Nazari, and E. A. J. I. C. L. Gharavol, "A novel DoS and DDoS attacks detection algorithm using ARIMA time series model and chaotic system in computer networks,"

IEEE Communications Letters, Vol. 20, No. 4,(2016), 700-703. DOI: 10.1109/LCOMM.2016.2517622

19. X. Chen, S. J. I. T. o. I. Yu, and Systems, "A collaborative intrusion detection system against DDoS for SDN," ***IEICE Transactions on Information and Systems,*** Vol. 99, No. 9, 2395-2399, 2016. DOI: 10.1587/transinf.2016EDL8016
20. Singh, Maninder Pal, and Abhinav Bhandari. "New-flow based DDoS attacks in SDN: Taxonomy, rationales, and research challenges." ***Computer Communications,*** (2020). https://doi.org/10.1016/j.comcom.2020.02.085
21. Q. Niyaz, W. Sun, and A. Y. J. a. p. a. Javaid, "A deep learning based DDoS detection system in software-defined networking (SDN)," arXiv preprint arXiv: *1611.07400* 2016. DOI: 10.4108/eai.28-12-2017.153515
22. D. Arivudainambi, V. K. KA, S. S. J. N. C. Chakkaravarthy, and Applications, "LION IDS: A meta-heuristics approach to detect DDoS attacks against Software-Defined Networks," ***Neural Computing and Applications*** Vol. 31, No. 5, (2019), 1491-1501. https://doi.org/10.1007/s00521-018-3383-7
23. M. Yazdani, F. J. J. o. c. d. Jolai, and engineering, "Lion optimization algorithm (LOA): a nature-inspired metaheuristic algorithm," ***Journal of Computational Design and Engineering***, Vol. 3, No. 1, (2016), 24-36. https://doi.org/10.1016/j.jcde.2015.06.003
24. I. Sreeram, V. P. K. J. A. c. Vuppala, and informatics, "HTTP flood attack detection in application layer using machine learning metrics and bio inspired bat algorithm," ***Applied Computing and Informatics***, Vol. 15, No. 1 (2019): 59-66.https://doi.org/10.1016/j.aci.2017.10.003
25. N. Dayal, P. Maity, S. Srivastava, R. J. S. Khondoker, and C. Networks, "Research trends in security and DDoS in SDN," ***Security and Communication Networks,*** Vol. 9, No. 18, (2016), 6386-6411. https://doi.org/10.1002/sec.1759
26. Bala, Ritu, and Ritu Nagpal. "A REVIEW ON KDD CUP99 AND NSL-KDD DATASET." International ***Journal of Advanced Research in Computer Science,*** Vol. 10, No. 2 (2019), 64.

https://doi.org/10.26483/ijarcs.v10i2.6395

27. I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in ***ICISSP***, (2018), 108-116. DOI: 10.5220/0006639801080116
28. S. Alelyani, J. Tang, and H. Liu, "Feature selection for clustering: A review," in *Data Clustering*: Chapman and Hall/CRC,(2018),29-60. https://doi.org/10.1007/springerreference_63708
29. A. B. Brahim, M.. Limam, , "Ensemble feature selection for high dimensional data: a new method and a comparative study," ***Advances in Data Analysis and Classification,*** Vol. 12, No. 4, (2018), 937-952. https://doi.org/10.1007/s11634-017-0285-y
30. T. J. Bihl, K. W. Bauer, M. A. J. I. T. o. I. F. Temple, "Feature selection for RF fingerprinting with multiple discriminant analysis and using ZigBee device emissions," ***IEEE Transactions on Information Forensics and Security,*** Vol. 11, No. 8, (2016), 1862-1874. DOI: 10.1109/TIFS.2016.2561902.
31. S. Saremi, S. Mirjalili, and A. J. A. i. E. S. Lewis, "Grasshopper optimisation algorithm: theory and application," ***Advances in Engineering Software***, Vol. 105, (2017), 30-47. https://doi.org/10.1016/j.advengsoft.2017.01.004
32. S. Saremi, S. Mirjalili, S. Mirjalili, and J. S. Dong, "Grasshopper Optimization Algorithm: Theory, Literature Review, and Application in Hand Posture Estimation," in *Nature-Inspired Optimizers*: Springer, (2020), 107-122. https://doi.org/10.1007/978-3-030-12127-3_7
33. Abualigah, Laith, and Ali Diabat. "A comprehensive survey of the Grasshopper optimization algorithm: results, variants, and applications." ***Neural Computing and Applications,*** (2020), 1-24. https://doi.org/10.1007/s00521-020-04789-8
34. Bansal, Priti, Sachin Kumar, Sagar Pasrija, and Sachin Singh. "A hybrid grasshopper and new cat swarm optimization algorithm for feature selection and optimization of multi-layer perceptron." ***Soft Computing,*** (2020), 1-27. https://doi.org/10.1007/s00500-020-04877-w
35. J. Luo, H. Chen, Y. Xu, H. Huang, and X. J. A. M. M. Zhao, "An improved grasshopper optimization algorithm with application to financial stress prediction," ***Applied Mathematical Modelling***, Vol. 64, (2018), 654-668. https://doi.org/10.1016/j.apm.2018.07.044
36. Joshi, Susheel Kumar, and Jagdish Chand Bansal. "Parameter tuning for meta-heuristics." ***Knowledge-Based Systems*** 189 (2020), 105094. https://doi.org/10.1016/j.knosys.2019.105094
37. H. Binol, I. Guvenc, E. Bulut, and K. J. E. L. Akkaya, "Hybrid evolutionary search method for complex function optimisation problems," ***Electronics Letters,*** Vol. 54, No. 24, (2018), 1377-1379. DOI: 10.1049/el.2018.6506
38. D. Ackley, *A connectionist machine for genetic hillclimbing*. Springer Science & Business Media, (2012). https://doi.org/10.1007/978-1-4613-1997-9
39. M. Babagoli, M. P. Aghababa, and V. J. S. C. Solouk, "Heuristic nonlinear regression strategy for detecting phishing websites," ***Soft Computing,*** Vol. 23, No. 12, (2019), 4315-4327. https://doi.org/10.1007/s00500-018-3084-2
40. S. Mirjalili and A. J. A. i. e. s. Lewis, "The whale optimization algorithm," ***Advances in engineering software,*** Vol. 95, (2016), 51-67. https://doi.org/10.1016/j.advengsoft.2016.01.008
41. Haddadi, Mohamed, and Rachid Beghdad. "A Confidence Interval Based Filtering Against DDoS Attack in Cloud Environment: A Confidence Interval Against DDoS Attack in the Cloud." ***International Journal of Information Security and Privacy (IJISP)***, Vol. 14, No. 4, (2020), 42-56. DOI: 10.4018/IJISP.2020100103
42. S. Bista and R. J. J. o. I. S. Chitrakar, "DDoS Attack Detection Using Heuristics Clustering Algorithm and Naïve Bayes Classification," ***Journal of Information Security,*** Vol. 9, No. 01, (2017), 33. DOI: 10.4236/jis.2018.91004

Persian Abstract

چکیده

امنیت سایبری به دلیل گسترش تهدیدات سایبری و آزار و اذیت‌های اینترنتی به محیطی وحشیانه و شرورانه تبدیل شده‌است. حمله انکار سرویس توزیع شده (DDoS) تهدیدی در شبکه است که منابع قربانیان را به خطر می اندازد. با توجه به نقش قابل توجه الگوریتم های بهینه سازی در شناسایی بسیار دقیق، قابلیت انطباق با حملات شبکه و نرخ هشدار کاذب قابل قبول، مطالعه حاضر یک روش ترکیبی مبتنی بر الگوریتم بهینه سازی ملخ اصلاح شده و الگوریتم ژنتیک (HMGOGA) برای شناسایی و جلوگیری از حملات DDoS پیشنهاد کرده است. HMGOGA بر معایب الگوریتم GOA سنتی از جمله سرعت همگرایی کم و گیر افتادن در بهینه محلی غلبه می کند. در این مقاله ، از الگوریتم پیشنهادی برای شناسایی حملات DDoS از طریق شبیه سازی مدل رگرسیون غیرخطی (NR) استفاده شده است. به منظور دستیابی به این منظور ، در ابتدا مهمترین ویژگیهای ترافیک شبکه با استفاده از روش جنگل تصادفی (RF) استخراج می شود. با حذف 55 ویژگی بی ربط از مجموع 77 ویژگی، ویژگی‌های منتخب نقش اساسی در عملکرد مدل پیشنهادی ایفا کرده اند. برای صحت سنجی کارایی ، همبستگی زیاد ویژگی های انتخاب شده با درخت تصمیم (DT) اندازه گیری شده است. متعاقباً ، HMGOGA با دو تابع هزینه معتبر و یک تابع هزینه پیشنهادی دیگر آموزش داده می شود که به آن امکان می دهد ترافیک مخرب را به درستی تشخیص دهد. جهت اعتبار سنجی و قابل استفاده بودن مدل پیشنهادی با دو تابع هزینه معتبر (عملکرد Sphere و Ackley) مقایسه می شود. نتایج تجربی ثابت کرده است که HMGOGA مبتنی بر NR-sigmoid به ترتیب با 99.90٪ و 99.34٪ دقت آموزش و آزمون، به ترتیب از سایر مدلهای پیاده سازی شده و روشهای GOA معمولی عملکرد بهتری داشته است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Routing Protocols for IoT Applications based on Distributed Learning

R. S. Raza[a,b], M. S. Al-Ani*[b]

[a] Technical College of Informatics, Sulaimani Polytechnic University, Sulaimani, Iraq
[b] College of Science and Technology, University of Human Development, Sulaymaniyah, Iraq

P A P E R I N F O

Paper history:
Received 29 January 2021
Received in revised form 19 February 2021
Accepted 02 March 2021

Keywords:
Contextual
Internet of Things
Protocol
Learning
Routing

A B S T R A C T

The IPv6 routing protocol for lossy and low power networks (RPL) was introduced in March 2012 by the Internet Engineering Task Force (IETF) as the standard routing protocol for the Internet of Things (IoT). Since that time, it has had various applications in IoT. Despite meeting the IoT network necessities by RPL, some unanswered issues have not been devised primarily for IoT usages. However, gathering a large amount of data from these networks with videos and images typically leads to traffic congestion in the network's central part. For providing a solution for this issue, the content-centric routing (CCR)-based RPL is proposed in the present study, where the content specifies the routing pathways. It is possible to attain a larger data aggregation ratio by routing the relevant data to the middle relaying nodes for the process. Thus, effective traffic is generated in the network. Subsequently, latency is significantly reduced. Moreover, energy use is decreased on wireless communication. More integration was conducted between IETF RPL protocol and CCR, using the MATLAB platform. Finally, according to simulated and implemented results, the CCR-based RPL behavior based on the high packet transfer rates is improved, and the numbers of dead nodes are reduced. High energy efficiency and low delay rates are obtained in communication using the proposed routing method.

***doi**: 10.5829/ije.2021.34.04a.08*

1. INTRODUCTION

Recently, there has been a high interest in the distributed computing in wireless networks, especially in the emerging pattern of Internet of Things (IoT) communications, where IoT instruments are supplied with independent processes, communication and storage capacities [1]. The important consequences include (i) a more economical approach with locally decreased data volume by in-network processing, (ii) just the processed outcomes are routed rather than all raw data, directed across a costly (multi-hop) wireless network. Therefore, bandwidth and energy are saved, the latency is decreased, and the network life is prolonged in an IoT network with restricted resources [2].

The IoT influences various application domains, including smart grids [3, 4], home automation [5], drone-based systems [6], healthcare systems [7], and industrial supervision [8, 9]. It is essential to use new wireless technologies to support innovative IoT applications and allowing large-scale connectivity among IoT policies. However, the combination of the IoT ecosystem and current wireless networks results in many issues, including self-organizing operation, coincidence with human-kind systems, and limited communication sources [7, 10]. Besides, the IoT devices are often machine-type ones, which have a significant difference from conventional human-type tools, such as smartphones based on performance needs, memory, computation, traffic patterns, and energy constraints [11]. Additionally, the IoT tools require short packets, ultrareliable transmissions, and low latency. Thus, re-planning of the current wireless networks is required for solving these IoT challenges.

A distributed learning approach with the Low Power Networks (RPL) protocol (CCR-based RPL protocol) was presented by Azari and Cavdar [12] to solve this problem to collect data in multi-hop IoT networks effectively as represented in Figure 1. Appropriate allocation of restricted resources is achieved using the

*Corresponding Author Institutional Email: *muzhir.al-ani@uhd.edu.iq* (M. S. Al-Ani)

Please cite this article as: R. S. Raza, M. S. Al-Ani, Routing Protocols for IoT Applications based on Distributed Learning, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 825-831

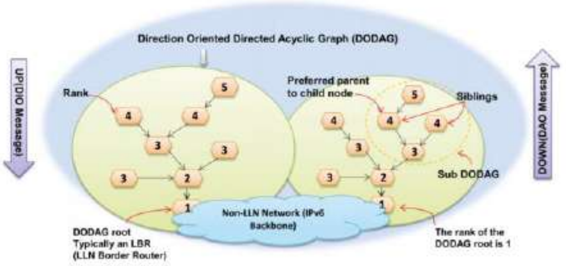


Figure 1. The architecture of the RPL routing domain

distributed method concerning the IoT tools' Quality of Service (QoS) needs [13]. These tools should have the capability to separately obtain their communication resources as supposing their regular communication with the root node is not practical, given their severe resource constraints. Moreover, the root in a massive IoT cannot manage the communication resources in all tools promptly. Thus, the distribution of resources is essential for all assignment occasions in the IoTs, considering the limited capacities of the IoT tools in terms of computation and memory. Deployment of the IoT over the available networks with limited communication resources would be confined by a resource assignment framework that meets the pre-described requirements.

The current research focuses on two critical dimensions of IoT networks: energy efficiency and transfer rates in a network having a high and heavy dynamic load [14]. A new objective functions (OF) is presented in this study, which periodically investigates the last states of node chains in the pathway to the root while avoiding the passage of additional message requests in the network. Additionally, this work provides a new routing metric to respond to the network's dynamicity requirement. Besides, a new approach is presented to parent selection.

2. REVIEW OF LITERATURE

Various research works have been conducted on the adaptive routings in objective functions and IoT, including energy-consumption, load balancing, and time delay. However, there are a rare number of studies conducted on distributed learning in routing. Some studies related to context information-based routing are presented in this section.

Mohamed et al. [5] studied the proactive routing protocols with the highest energy efficiency for a homogeneous system. The Wireless Sensor Network (WSN) energy efficiency results from three fundamental necessities: network overhead within network arrangement and re-arrangement, selecting a route for data transmission, and fault tolerance or network adaptability. For guaranteeing a longer network lifetime, it is required to have a changeable network, the overhead must be confined to a minimum rate, and multi-hop routing must be used. Thus, selecting the energy-efficient routing protocols must be done cautiously based on application necessities and system requirements. This work attempts to deal with the network lifetime analysis for the reactive routing protocols in heterogeneous and homogeneous systems.

Aman et al. [15] investigated the RPL acceptance in the Cooja simulator by comparing the real networks' findings. They showed higher energy consumption by both the sink and client nodes in a real network compared to the Cooja simulation calculation. The packet delivery rate (PDR) was found to be lower in actual scenarios and is considerably decreased by the network distance and size. However, interference is not involved from external wireless, signal reflection, and asymmetric links. This model is yet extensively employed in different works.

A new uplink bandwidth assignment algorithm was presented by Mardani et al. [9] for highly dense 4G networks of the human-to-human (H2H)/ machine-to-machine (M2M) services in a co-existing setting. The algorithm utilized interval type-2 fuzzy logic for dealing with system ambiguities. The intelligent type-2 fuzzy algorithm specifies the optimum bandwidth ratio for every M2M/ H2H service flow. To determine the correct bandwidth ratio for every type of flow, the solution used the M2M/H2H necessities, including the Human Type Communication (HTC) output, the power level of Machine Type Communications (MTC) tools for non-real-time services, and the highest potential for real-time services. Assessment of the suggested system performance was done based on throughput, bandwidth utilization, and delay.

A context awareness framework known as Awareness Cognition (AC) was proposed by Kim and Yoon [16] that focused on the Ambient Intelligence (AmI) system's middleware layer. Using this framework, the prediction problem was solved by integrating multiple contexts for detecting personalized knowledge. Additionally, utilizing knowledge-based cognition approaches for discovering future changes, a fusion contextual learning model was generated for behavioral knowledge detection. Besides, modeling was provided for collecting data from various sensors in the AmI system in a single meaningful context [13, 17].

Two algorithms with two constraints were presented by Huang et al. [18] for challenging the multicast routing problem for multimedia communication in the IoT. Multiple constraints are integrated into a complete metric using the entropy method. Thus, the proposed algorithms significantly decreased the complication of multi-constrained multicast routing problems, and some popular algorithms can be applied for problem-solving. Additionally, the theoretical analysis was suggested on the approximation and complexity of the suggested algorithms, and wide simulations were carried out to assess the algorithm's performance. According to analytical and experimental findings, one of the proposed algorithms is better than a representative multi-constrained multicast routing algorithm in terms of speed and accuracy. The results provide an in-depth perception of the multicast routing algorithm design for multimedia communications in IoT.

Rani et al. [19] proposed an improved solution for organizing objects to implement an energy-efficient and scalable IoT. Firstly, the framework was presented to deploy the IoT with scalability features providing higher extensibility. Afterward, considering the framework, an optimization outline can support the deployment of an IoT with energy efficiency. This optimization outline is confined by the loads on wireless links and energy expenses. Compared to conventional WSN outlines in terms of network lifetime, time, and scalability, various numerical tests confirm superiority of the proposed outline. This work addresses the challenges of the way of using heterogeneity advantages. Nevertheless, improvement of end-to-end delay, throughput parameters, packet delivery ratios, and data compression approaches are proposed to achieve more effective green IoT.

An effective self-organization protocol called ETSP was presented by Qui et al. [20] for sensor networks of IoTs, saving energy and allowing a prolonged lifetime for a network by restricting a tree-based network. The nodes' weight, such as hop, residual energy, distance within the nodes, and a number of child nodes were used to determine whether it can be a sink node. Therefore, the depth of the tree is improved by ETSP. The topology of the network is drastically changed over the data transmission procedure. Because the sink nodes consume energy faster than other nodes, each sink node is dynamically re-chosen. The simulation results showed that with ETSP, reliable tree-based networks could be generated, the energy consumption can be reduced, and the sensor network's lifetime can be extended.

Shen et al. [21] presented an energy-efficient centroid-based routing protocol (EECRP) for controlling the WSN-assisted IoT energy. To this end, they solved the forming clusters problem considering the distance to the energy centroid. An enhancement algorithm was presented in the present study regarding the number of cluster head nodes and the number of dead nodes. The simulation results indicated that with the base station (BS) deployment in the network, it would be possible to transfer a great amount of data by the EECRP with very low energy dissipation. Besides, the EECRP has a longer network lifetime than the GEEC, LEACH-C, and LEACH. This protocol will be enhanced for future study by identifying the multi-hop pathway from cluster head (CH) nodes to BS. For transferring data packets, a multi-hop pathway is used by the CH nodes. It is hoped that the future protocols have good performance for the BS outside the network.

3. STATEMENT OF THE PROBLEM

RPL, which is primarily devised for low-power lossy networks, possesses numerous notable characteristics, including self-healing mechanism, loop-freeness, low-battery use, and fast topology construction. Nevertheless, it fails to address issues of a network with a high traffic rate as it was chiefly devised for low traffic networks. When there is a high network traffic, RPL cannot control it well, making various difficulties for the network, such as energy depletion, load imbalance, and high packet loss rate. It would be more troublesome when a depleted node is the only intermediate node for a network section close to the root. RPL problems are classified under heavy and highly dynamic loads as follows:

1- With the computation of the rank through two OFs including Objective Function Zero (OF0) and minimum rank with hysteresis objective function (MRHOF) in normal RPL, many studies have been conducted for changing the RPL objective functions based on related research works. Nevertheless, as it is known, the former parents of a node in the sequence are not considered by the ordinary RPL OFs and other proposed OFs. A node could seem proper for serving as a parent; however, the node parent or another parent in a parent sequence could have low remaining power or small buffer space, which causes inappropriate selection of the parent in the network with high traffic.

2- The route is created concerning the rank. When attempting to connect the network or alter its parent by a node, the network is selected with a lower rank value. Nevertheless, the parents' rank is obtained in the primary phases of connecting the network. Various difficulties could appear in the pathway to the route followed by the rank computation. Thus, the rank value in a high traffic dynamic network cannot completely represent the last and real mode of the candidate parents.

4. METHODOLOGY

As a node connects to a Destination Oriented Directed Acyclic Graph (DODAG) and if a DODAG Information Object (DIO) message is received by it, the node can process it via the following ways:

1) Removing the DIO package for some RPL criteria.
2) Processing the message for keeping its location in the network
3) Improving its location by gaining a lower rank in DODAG.

When a node's rank decreases, the node must eliminate all parents with a rank lower than its new rank from its parents' list. Thus, the development of a loop in the network is prevented. Following this phase, the nodes each has a default route to the root and they would be able to send their data packages to the root.

According to the simulation results, the efficiency of the CCR-based RPL routing approach in data aggregation and recognition in IOT network is high. In this section, using the content-based RPL approach in the direction of data aggregation rate, we compare the number of live nodes, energy consumption, the network balance, and correct data transfers with those in the conventional RPL approach.

Similar to the properties given in Table 1, the simulation's completion and the node routings in the network can be observed.

For increasing the routing efficiency in IOT network, we presented a content-based system in this work. In the presented system, each node utilizes the root-inquiry packages for recognizing the parent nodes to become aware of the number of packages arriving in the root. Utilizing this awareness, the level of confidence is calculated for the route that the parent provides. In cases of malignancy of the parent and when the level of confidence for parent node calculation is lower, the respective node chooses a parent with a higher confidence level among the candidate parents. Using this method, each node can efficiently prevent malignant nodes. It is possible to gain a higher data collection ratio by routing the intermediate relay node data in the processing order. Thus, it is effective in the reduction of the network traffic rate. Consequently, the delay in the data transfer can be significantly reduced. Moreover, we can annihilate data transfers following data collection, resulting in a decrease in energy consumption in wireless communications and saving battery energy consumption.

Three major units are included in this method:

1) Data collection unit: It gains information about the number of data packages arriving at the root by periodically sending the root-inquiry packages.
2) Confidence level calculation unit: This unit uses the information presented by the data collection unit and computes the confidence level of the route presented by the parent using the following equation:

$$T(p) = 1 - \frac{pk\ sent_{ij} - pk\ delivered_{ij}}{pk\ sent_{ij}}$$

$$if\ pkrecieved = pk_{delivered} \qquad (1)$$

$$T(c) = T(c) - 0.01 \qquad else$$

where *T(p)*: Confidence level of each parents (neighbors), *Pk sendij*: Total number of packets sent to the root by specific node, $pk_{delivered}$: Total number of packets arrived to the root through each parent (neighbors)

3) Parent selection unit: It chooses a parent with a higher level of confidence among the candidate parents as the selected parent.

TABLE 1. Properties of the simulation employed in our work.

Parameter	Value
Network area	$800m \times 800m$
number of nodes	50
Velocity of Node	0
Transfer range	100 m
Load size	512 Bytes
Packet transfer rate	25 KB
Data aggregation rate	0.0011 µJ
Data reception energy consumption	8.22 µJ
Data transfer energy consumption	9.72 µJ
Simulation numbers	20

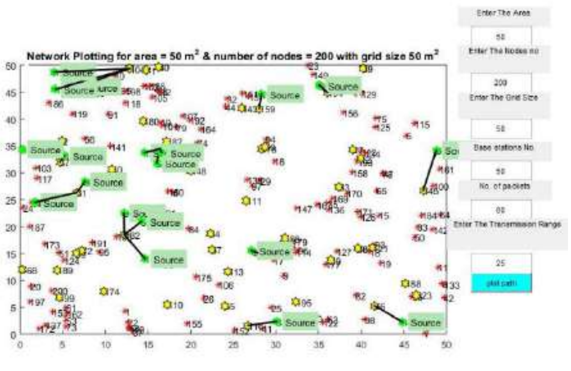


Figure 2. Routing the network topology

5. EVALUATION

To assess the protocol proposed in our work, the CCR-based RPL was simulated using the MATLAB simulator, commonly used for IoT. Then, the CCR-RPL was compared with conventional RPL. MRHOF was adjusted as OF of ordinary RPL, and the protocol factors were tuned based on the CCR-based RPL factors (Table 1). There are 800 nodes in our setting with 20 BS, which were established in 50m × 50m. The BS serves as the root, and the transmission range is 25m.

In this setting, duty cycling was disabled for reaching a high load in the network, and the first-in-first-out (FIFO) line was utilized with a volume of 80 packets. The simulation setting and the simulation parameters are represented in Figure 2 and Table 1. Besides, various traffic rates and nodes are considered for assessing the protocol proposed in our work under different conditions.

Firstly, a comparison was conducted based on the queue loss ratio. Figure 3 indicates two protocols with varying traffic loads (rising). According to the results, the numbers of the live nodes are increased by the network's CCR-based RPL. Nevertheless, the other important property shown in Figure 3 indicates the worst case of the queue loss ratio in the nodes in varying traffic loads. We interpret it as the CCR-based RPL capability to make a more uniform DODAG concerning the network load.

Following application of the simulation, the following results were obtained:

Figures 3-8 indicate the higher proficiency of the CCR-RPL approach than the ordinary RPL in increasing system efficiency and reducing energy consumption. Thus, it reduces the number of dead nodes and energy consumption in the system and increases the system efficiency and power.

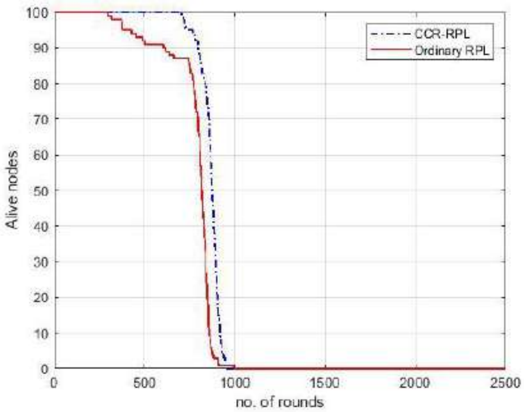


Figure 3. Comparing the number of the live nodes

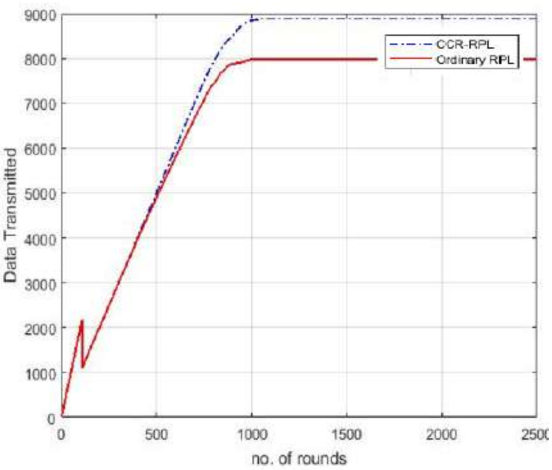


Figure 4. Comparing the rates of transferred packages

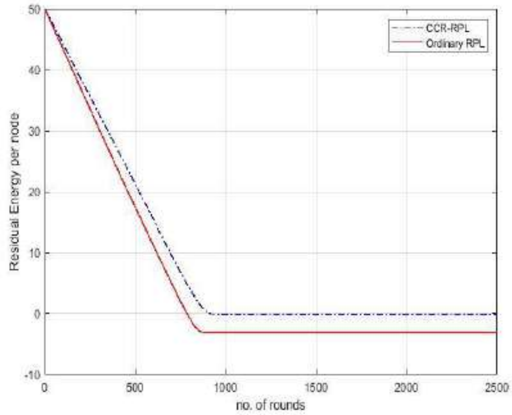


Figure 5. Comparing the rates of the remaining energy

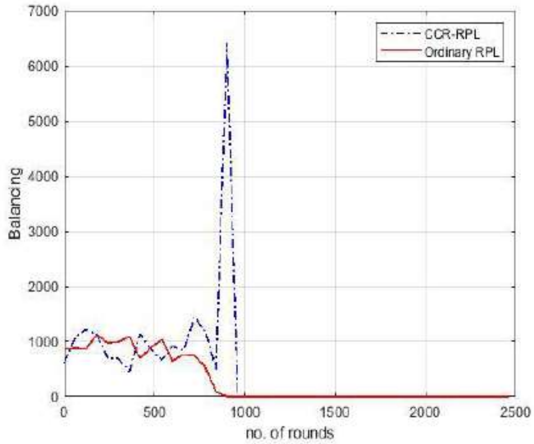


Figure 6. Comparing the balances dominating the system

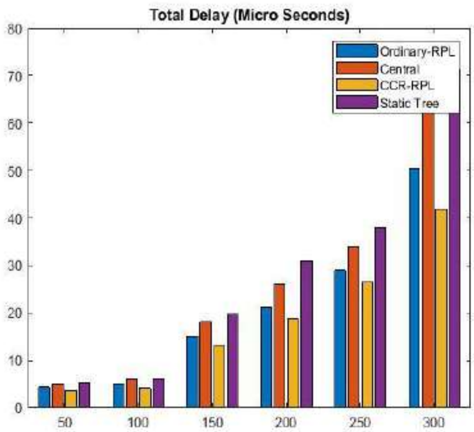


Figure7. The delay rates in package transfers

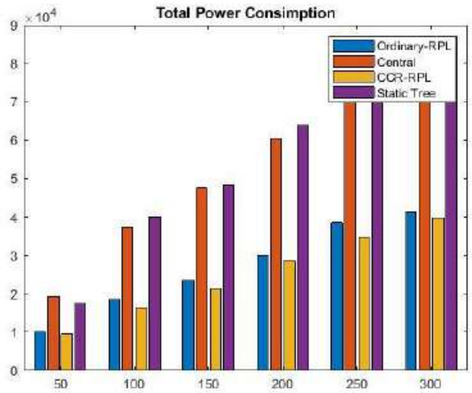


Figure 8. Comparing the rates of power consumed

6. CONCLUSION

RPL routing protocol problems with heavy and dynamic load were addressed in the present work, focusing on the network lifetime and energy. It was observed that an ordinary RPL could not provide effective control over the heavy and dynamic loads. To provide a solution for this challenge, a load balancing and context-aware protocol was proposed considering the parent-chain rank before selecting the chain's ultimate parent as the node's chosen to parent. Thus, the load was tried to be balanced in the network. The present study took into account the residual queue and energy level of the candidate parents.

Additionally, a proper parent's rush was prevented, making instability problems and a high control message rate for the network. Our proposed protocol was assessed in MATLAB in varying conditions. It was proved that the performance of CCR-RPL is significantly better than RPL, while it does not impress a high overload for the network.

7. REFERENCES

1. Xie, J., Chen, W., Dai, H., Liu, S. and Ai, W., "A distributed cooperative learning algorithm based on zero-gradient-sum strategy using radial basis function network", ***Neurocomputing***, Vol. 323, (2019), 244-255. https://doi.org/10.1016/j.neucom.2018.10.001
2. Debroy, S., Samanta, P., Bashir, A. and Chatterjee, M., "Speed-iot: Spectrum aware energy efficient routing for device-to-device iot communication", ***Future Generation Computer Systems***, Vol. 93, (2019), 833-848. https://doi.org/10.1016/j.future.2018.01.002
3. Thivakaran, T. and Sakthivel, T., "Guard: An intrusion detection framework for routing protocols in multi-hop wireless networks", ***Wireless Networks***, Vol. 25, No. 2, (2019), 819-836. https://doi.org/10.1007/s11276-017-1594-y
4. Nielsen, J.J., Madueño, G.C., Pratas, N.K., Sørensen, R.B., Stefanovic, C. and Popovski, P., "What can wireless cellular technologies do about the upcoming smart metering traffic?", ***IEEE Communications Magazine***, Vol. 53, No. 9, (2015), 41-47. https://doi.org/10.1109/MCOM.2015.7263371
5. Mohamed, R.E., Saleh, A.I., Abdelrazzak, M. and Samra, A.S., "Survey on wireless sensor network applications and energy efficient routing protocols", ***Wireless Personal Communications***, Vol. 101, No. 2, (2018), 1019-1055. https://doi.org/10.1007/s11277-018-5747-9
6. Mozaffari, M., Saad, W., Bennis, M. and Debbah, M., "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs", ***IEEE Transactions on Wireless Communications***, Vol. 15, No. 6, (2016), 3949-3963. https://doi.org/10.1109/TWC.2016.2531652
7. Hwang, K. and Chen, M., "Big-data analytics for cloud, iot and cognitive computing, John Wiley & Sons, (2017).
8. Naguib, A., Saad, W. and Shokair, M., "Remaining energy aware ml-csma/tdma hybrid mac protocol for lte-m2m wireless network", in 2019 International Conference on Innovative Trends in Computer Engineering (ITCE), IEEE. 322-327.
9. Mardani, M.R., Mohebi, S. and Ghanbari, M., "Energy and latency-aware scheduling under channel uncertainties in lte networks for massive iot", ***Wireless Personal Communications***, Vol. 103, No. 3, (2018), 2137-2154. https://doi.org/10.1007/s11277-018-5901-4
10. Park, T. and Saad, W., "Distributed learning for low latency machine type communication in a massive internet of things", ***IEEE Internet of Things Journal***, Vol. 6, No. 3, (2019), 5562-5576. https://doi.org/10.1109/JIOT.2019.2903832
11. Dawy, Z., Saad, W., Ghosh, A., Andrews, J.G. and Yaacoub, E., "Toward massive machine type cellular communications", ***IEEE Wireless Communications***, Vol. 24, No. 1, (2016), 120-128. https://doi.org/10.1109/MWC.2016.1500284WC
12. Azari, A. and Cavdar, C., "Self-organized low-power iot networks: A distributed learning approach", in 2018 IEEE Global Communications Conference (GLOBECOM), IEEE., 1-7.
13. Feizi, A., "Convolutional gating network for object tracking", ***International Journal of Engineering***, Vol. 32, No. 7, (2019), 931-939. https://doi.org/10.5829/ije.2019.32.07a.05
14. Hossein Motlagh, N., Mohammadrezaei, M., Hunt, J. and Zakeri, B., "Internet of things (iot) and the energy sector", ***Energies***, Vol. 13, No. 2, (2020), 494. https://doi.org/10.3390/en13020494
15. Aman, M.S., Yelamarthi, K. and Abdelgawad, A., "A comparative analysis of simulation and experimental results on rpl performance", in 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), IEEE., 483-487.
16. Kim, S. and Yoon, Y.-I., "Ambient intelligence middleware architecture based on awareness-cognition framework", ***Journal Of Ambient Intelligence And Humanized Computing***, Vol. 9, No. 4, (2018), 1131-1139. https://doi.org/10.1007/s12652-017-0647-5
17. Firouzian, I., Zahedi, M. and Hassanpour, H., "Cycle time optimization of processes using an entropy-based learning for task allocation", ***International Journal of Engineering***, ***Transactions B: Applications,*** Vol. 32, No. 8, (2019), 1090-1100. https://doi.org/10.5829/ije.2019.32.08b.05
18. Huang, J., Duan, Q., Zhao, Y., Zheng, Z. and Wang, W., "Multicast routing for multimedia communications in the internet of things", ***IEEE Internet of Things Journal***, Vol. 4, No. 1, (2016), 215-224. https://doi.org/10.1109/JIOT.2016.2642643
19. Rani, S., Talwar, R., Malhotra, J., Ahmed, S.H., Sarkar, M. and Song, H., "A novel scheme for an energy efficient internet of things based on wireless sensor networks", ***Sensors***, Vol. 15, No. 11, (2015), 28603-28626. https://doi.org/10.3390/s151128603
20. Qiu, T., Liu, X., Feng, L., Zhou, Y. and Zheng, K., "An efficient tree-based self-organizing protocol for internet of things", ***Ieee Access***, Vol. 4, (2016), 3535-3546. https://doi.org/10.1109/ACCESS.2016.2578298

21. Shen, J., Wang, A., Wang, C., Hung, P.C. and Lai, C.-F., "An efficient centroid-based routing protocol for energy management in wsn-assisted iot", ***Ieee Access***, Vol. 5, (2017), 18469-18479. https://doi.org/10.1109/ACCESS.2017.2749606

Persian Abstract

چکیده

در مارس 2012 پروتکل مسیریابی IPv6 برای شبکه های کم مصرف (RPL) توسط Task Force مهندسی اینترنت به عنوان پروتکل استاندارد مسیریابی برای اینترنت اشیا (IoT) تأیید شد. از آن زمان ، این برنامه کاربردهای مختلفی در اینترنت اشیا دارد. علی رغم تأمین نیازهای شبکه اینترنت اشیا توسط RPL ، برخی از موضوعات بی پاسخ برای کاربردهای اینترنت اشیا مطرح نشده است. با این حال ، جمع آوری مقدار زیادی از داده ها از این شبکه ها با فیلم و تصاویر به طور معمول منجر به ازدحام ترافیک در قسمت مرکزی شبکه می شود. برای این موضوع، در این مقاله روش مسیریابی محتوا محور RPL مبتنی بر CCR ارایه شده است. در این روش محتوا مسیرهای مسیریابی را مشخص می کند. با مسیریابی داده های مربوطه به گره های رله میانی برای فرآیند ، می توان به یک نسبت تجمع داده بزرگتر دست یافت. بنابراین ، ترافیک موثر در شبکه ایجاد می شود. پس از آن ، تاخیر به طور قابل توجهی کاهش می یابد. علاوه بر این ، استفاده از انرژی در ارتباطات بی سیم کاهش می یابد. ادغام بیشتر بین پروتکل IETF RPL و CCR با استفاده از سیستم عامل MATLAB انجام شد. سرانجام ، با توجه به نتایج شبیه سازی شده و پیاده سازی شده ، رفتار RPL مبتنی بر CCR بر اساس نرخ بالای انتقال بسته بهتر است و تعداد گره های مرده کاهش می یابد. بازده انرژی بالا و تسریع در انتقال حاصل می شود.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 832-842

International Journal of Engineering

Journal Homepage: www.ije.ir

A Game-theoretic Approach for Robust Federated Learning

E. Tahanian*, M. Amouei, H. Fateh, M. Rezvani

Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

PAPER INFO

Paper history:
Received 20 November 2020
Received in revised form 28 December 2020
Accepted 14 January 2021

Keywords:
Federated Learning
Game Theory
Byzantine Model
Adaptive Averaging

ABSTRACT

Federated learning enables aggregating models trained over a large number of clients by sending the models to a central server, while data privacy is preserved since only the models are sent. Federated learning techniques are considerably vulnerable to poisoning attacks. In this paper, we explore the threat of poisoning attacks and introduce a game-based robust federated averaging algorithm to detect and discard bad updates provided by the clients. We model the aggregating process with a mixed-strategy game that is played between the server and each client. The valid actions of the clients are to send good or bad updates while the server can accept or ignore these updates as its valid actions. By employing the Nash Equilibrium property, the server determines the probability of providing good updates by each client. The experimental results show that our proposed game-based aggregation algorithm is significantly more robust to faulty and noisy clients in comparison with the most recently presented methods. According to these results, our algorithm converges after a maximum of 30 iterations and can detect 100% of the bad clients for all the investigated scenarios. In addition, the accuracy of the proposed algorithm is at least 15.8% and 2.3% better than the state of the art for flipping and noisy scenarios, respectively.

doi*: 10.5829/ije.2021.34.04a.09*

1. INTRODUCTION

As datasets grow, the optimization of learning model parameters needs distribution across multiple machines. The idea of federated learning has recently been proposed, in which a shared global model is trained with the cooperation of a central server and some participants named clients [1-7], as can be seen in Figure 1. In other words, the clients train the model using their own local datasets and send it back to the central server. The server aggregates the information sent by the clients to update the shared global model. Afterward, the server sends the updated global model to some of the clients and this process is repeated again. Since the clients send only the model and not the data to the server, the data privacy is preserved.

One of the most important concerns about federated learning is sending bad updates by faulty or malicious clients. The researchers showed that only one bad client can compromise the model as well as the result in a convergence problem [4]. Thus, the researchers have tried to mitigate this problem by proposing different robust federated learning approaches [1, 8-10].

However, some of these techniques impose computational cost in comparison with the conventional averaging such as Federated Averaging (FA) [1], especially for a large number of clients. In addition, most of these techniques do not consider the number of data points used by each client to train the local models.

In this paper, we propose a Game-based robust Federated Averaging algorithm (GFA) to detect and discard bad updates provided by the clients. The proposed method uses an iterative averaging algorithm to highlight the effect of the good updates sent by the majority of the clients. At the end of this iterative algorithm, a trustworthiness is assigned to each client that can be used to put the client in one of the good or bad sets. Finally, the server considers the probability of providing good updates to the model by each client. These probabilities can be computed by considering a mixed-strategy game between the central server and each client that exists in the good client set. The valid actions of the clients are to

*Corresponding Author Institutional Email: e.tahanian@shahroodut.ac.ir (E. Tahanian)

Please cite this article as: E. Tahanian, M. Amouei, H. Fateh, M. Rezvani , A Game-theoretic Approach for Robust Federated Learning , International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 832-842

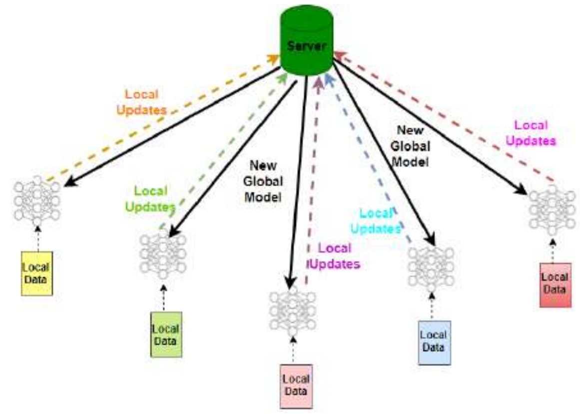


Figure 1. Architecture for a federated learning system with five benign clients that communicate with a central server periodically to learn a global model

send good or bad updates while the server can accept or ignore these updates. By employing the Nash Equilibrium property [11], the server determines the clients' probability to provide good updates to the model. We summarize our main contributions as follows:

- We propose an iterative averaging algorithm for the server to obtain the trustworthiness of each update and a robust estimate of the final model, simultaneously.
- We model the problem using a mixed-strategy game between the central server and each client.
- We apply the Nash Equilibrium property to compute the probability of providing good updates from each client.

We provide a thorough empirical evaluation of the effectiveness and efficiency of our proposed robust federated learning method. The results show that our method provides both higher accuracy and faster convergence than the existing methods. Specifically, our algorithm converges after a maximum of 30 iterations and can detect 100% of the bad clients for all the investigated scenarios. Furthermore, the accuracy of the proposed algorithm is at least 15.8% and 2.3% better than the state of the art for flipping and noisy scenarios, respectively.

The rest of this paper is organized as follows. The related work is discussed in Section 2. Section 3 describes the proposed federated learning as well as the aggregation algorithm. The experimental results are reported and discussed in Section 4, followed by the conclusion in Section 5.

2. RELATED WORK

Federated learning for the first time was implemented by Google to predict users' text input within a large number of mobile devices without sending private data [2, 12]. One of the main elements of Federated learning is the aggregation operator. Several federated aggregation operators have been presented in literature. FedAvg [13] updates the global model by averaging the parameters of the local models. This algorithm was used for recognizing out-of-vocabulary words [14] and improving the mobile keyboard prediction [15]. As a modification to FedAvg, Federated Stochastic Variance Reduced Gradient (FSVRG) [2] was presented to work with sparse data. In contrast to FedAvg and FSVRG, CO-OP [16] has been presented for asynchronous model updates. It merges any received client model with the global model. According to the difference in the age of the models, the local and global models merging is carried out using a weighting scheme, instead of directly averaging the models.

Due to the distributed scheme of federated learning, it is highly vulnerable to attacks against the learning models. As previously mentioned, sending bad updates by faulty or malicious clients is the most serious concern for federated learning. Consequently, the standard federated learning algorithms such as FedAvg [1, 13] are vulnerable to both model poisoning and data poisoning. To overcome this problem, researchers have proposed different robust averaging algorithms [8, 17, 18].

Some other researches have focused on vulnerability in federated learning known as a backdoor attack [3, 19]. In this kind of attack, the adversary tries to reduce the performance of the model on targeted tasks while maintaining good performance on the main task [19].

Authors in [8] proposed a byzantine-robust aggregation algorithm, referred to as KRUM, which is based on the similarity of the client updates. To solve the slow convergence problem of the KRUM, a faster algorithm known as MKRUM was introduced. Yin et al. [17] proposed a coordinate-wise median (COMED), a byzantine-robust statistical learning algorithm with a focus on statistical optimality.

Although the aforementioned researches take into account both model poisoning and data poisoning for a number of simple attack scenarios, the proposed methods can be computationally expensive when the number of clients is large. In contrast, the computational complexity of our method (GFA) can be shown to be considerably less. Moreover, the previous algorithms do not consider the number of data points used by each client to train the local models, while the GFA computes the averaging based on the dataset size of each client.

Moreover, our proposed method flexibly chooses the good clients based on the received information at each iteration. Therefore, it is more efficient in comparison with the works that use a pre-specified number of clients' information to update the global model, such as the work proposed by Xie et al. [18]. Especially, unlike the GFA, when all of the clients or the majority of them are good, the algorithm of Xie et al. [18] considers the pre-specified number of the clients as bad, which will affect the performance of the learning process.

Also, some works have focused on applying game theory to the federated learning system [20-22]. The authors in [20] proposed a contract theory-based incentive mechanism to motivate data owners that have high-quality local training data to join the learning processes. In [21], the authors presented the Stackelberg game model to analyze the transmission strategy and training data pricing strategy of the self-organized mobile device as well as the learning service subscription of the model owner in the cooperative federated learning system. Zou et al. [22] adopted an evolutionary game theory to model dynamic strategies of the mobile devices with bounded rationality in the federated learning system. Although all of these works used the game theory, none of them focused on the averaging algorithm in federated learning. In this paper, we use the game theory to propose a robust federated averaging algorithm to detect and discard bad updates provided by the clients.

3. FEDERATED LEARNING AND AGGREGATION ALGORITHM

In this section, we formulate the federated learning paradigm and propose our robust aggregation algorithm based on a game-theoretical approach.

3. 1. Federated Learning Model The main idea of federated learning is to perform the training of a deep neural network (DNN) using some clients by aggregating local models into a joint global model, as can be seen in Figure 1. Since the local training data never shared by the clients, the federated model can train on completely private data.

We suppose there are N clients where the ith client's dataset has n_i data points. At round t, the server randomly chooses a subset of clients (M_t) and sends them the latest global model (ω_t). Each client, for example, $i \in M_t$ that receives the model, updates it by training on its dataset, and derives a new local model (ω_{t+1}^i). Afterward, the chosen clients send back the new model to the central server. In this step, the server averages the received local models to achieve an updated global model according to Equation (1).

$$\omega_{t+1} = \sum_{i \subset M_t} \frac{n_i}{n} \omega_{t+1}^i, \tag{1}$$

where n is obtained as follows:

$$n = \sum_{i \subset M_t} n_i \; . \tag{2}$$

However, by using such an aggregation method, only a bad (malicious or faulty) client can lead to the wrong solution or prevent the DNN to be converged [4]. To solve this problem, in the next sections, we propose a novel aggregation algorithm based on the game theory in which the probability of providing good model updates by each client is considered. Table 1 contains a summary of the notations used in this paper.

3. 2. Adversary Model In this paper, we make the following assumptions regarding the adversary: (1) We assume that only less than half of the clients can be compromised; (2) the attacker controls the local training data of any compromised client; (3) it does not control the aggregation algorithm used by the server to average clients' updates and generate the new global model; (4) the attacker can not control the updates sent by the good clients and, (5) the data is distributed among the clients in an i.i.d fashion.

The adversary's goal in our work is to prevent the global model to converge. So, we propose a novel aggregation algorithm to overcome the convergence problem of the previous algorithms while the attacker follows the above-mentioned assumptions. In the rest of this paper, we use *bad* clients whenever we mean malicious or faulty clients.

3. 3. Aggregation Algorithm To implement robust federated learning, we should initially estimate the bad clients. To reach this goal, we propose a novel averaging algorithm as well as a game model in this Section.

3. 3. 1. Averaging Algorithm At each round, when the central server receives the local updates of the clients; it uses an adaptive averaging method. In this paper, we proposed to highlight the effect of the good updates sent by the majority of the clients [23, 24]. The proposed aggregation algorithm is detailed in Algorithm 1.

When the server receives the local updates of the clients, the server iteratively computes a weighed average as follows:

TABLE 1. The notation used in this paper.

Symbol	*Description*
N	The number of clients
n_i	The size of the ith client dataset
M_t	The number of clients that send updates at round t
m	The size of M_t
ω_{t+1}^i	The local model provided by the ith client at round t
ω_{t+1}	The global model sent to the clients
G_t	The set of clients considered good by the server
N_G	The size of G_t
AA_k	The weighed average of the received updates in kth iteration
y_i	The distance between ω_{t+1}^i and AA_{k-1}

$$AA_k = \frac{\sum_{i \in M_t} e^{-\alpha y_i} \omega_{t+1}^i}{\sum_{i \in M_t} e^{-\alpha y_i}}, \quad k > 1 \tag{3}$$

where AA_k is the weighted average of the received updates in *k*th iteration and y_i is the distance between ω_{t+1}^i and the weighted average at the previous iteration (AA_{k-1}), obtained as follows:

$$y_i = |\boldsymbol{\omega_{t+1}^i} - \mathbf{AA_{k-1}}|$$

It is to be noted that we compute the distance of each client's model with the current estimate of the aggregation values to estimate the trustworthiness of the clients. In other words, there is an inverse relationship between the trustworthiness of a client and the distance of its local model with the aggregated model obtained at each iteration of our adaptive aggregation algorithm. We employed an exponential decaying function to model such an inverse relationship as it shows promising results in our experiments. One can choose any other decaying function to compute the trustworthiness from the distance value of the models. In Equation (3), α is a constant parameter that controls the amount of *trustworthiness* ($e^{-\alpha y_i}$) considered for each update.

It should be noted that the iterative procedure starts with giving equal credibility to all clients, i.e., with an initial value of 1. Consequently, the initial aggregated model at the first iteration of the algorithm is calculated using a simple averaging as follows:

Algorithm 1 Robust Aggregation Algorithm

Require: $\boldsymbol{M_t}, \boldsymbol{n_i}, \boldsymbol{\omega_{t+1}^i}$, **K**, $\boldsymbol{\alpha}$

$\boldsymbol{G_t} \leftarrow \{\boldsymbol{i}: \boldsymbol{i} \in \boldsymbol{M_t}\}$

$\boldsymbol{m}$ **: the size of** $\boldsymbol{M_t}$

for k = 1, 2, 3,…, K do

  if k = 1 then

    $\mathbf{AA_1} = \frac{\sum_{i \in M_t} \omega_{t+1}^i}{m}$

  else

    for $\boldsymbol{i} \in \boldsymbol{M_t}$ **do**

      $\boldsymbol{y_i} = |\boldsymbol{\omega_{t+1}^i} - \mathbf{AA_{k-1}}|$

    end for

    Compute $\mathbf{AA_k}$ **according to Eq. (3)**

  end if

end for

Apply k-means to set of $\{e^{-\alpha y_1}, \dots, e^{-\alpha y_m}\}$ **and form two sets,** $\boldsymbol{G_t}$ **and** $\boldsymbol{B_t}$

Compute $\boldsymbol{p_t}^{\boldsymbol{i}}$ **using Game model**

$\boldsymbol{\omega_{t+1}} \leftarrow \frac{\sum_{i \in G_t} \boldsymbol{p_t}^{\boldsymbol{i}}\ \boldsymbol{n_i}\ \boldsymbol{\omega_{t+1}^i}}{\sum_{i \in G_t} \boldsymbol{p_t}^{\boldsymbol{i}}\ \boldsymbol{n_i}}$

return $\boldsymbol{\omega_{t+1}}$ **,** $\mathbf{G_t}$

$$AA_1 = \frac{\sum_{i \in M_t} \omega_{t+1}^i}{m} \tag{4}$$

where m is the size of M_t. By considering a stopping criterion (K), according to the variation of the *trustworthiness* of the clients, the iterative algorithm will be stopped. The main idea for our proposed aggregation algorithm is inspired by the iterative filtering algorithm proposed by [in the literature [25]. In this reference, a class of voting systems based on iterative filtering has been presented. In other words, in the first round, the simple average of the votes is calculated. Then, proportional to the inverse of the distance from the calculated average, an averaging weight is considered for each vote to compute a next round average. This process continues until the majority votes are close enough to each other and therefore the minority votes are filtered. The proof of convergence has been provided in the literature for two different discriminant functions [25]. We believe that a similar method can be used to prove the convergence of the iterative process in Algorithm 1. We leave this proof as an interesting idea for our future work.

Now, we expect the *trustworthiness* of the good clients are similar enough and spaced far enough from the bad clients [23]. So, the server can apply a one-dimensional *k-means* algorithm to put the clients in two separate clusters, *bad* clients (B_t) and *good* clients (G_t) with more than half of the clients. Thus, the server can average the updates by good clients regarding to the fraction of training data points provided by each client (Equation (5)).

However, due to the similarity of the updates provided by the clients, some of them may be misdiagnosed. Therefore, the server should consider a probability (${p_t}^i$) in the model aggregation algorithm for each client in G_t, according to Equation (5).

$$\omega_{t+1} = \frac{\sum_{i \in G_t} {p_t}^i\ n_i\ \omega_{t+1}^i}{\sum_{i \in G_t} {p_t}^i\ n_i} \tag{5}$$

The clients' probability to provide good updates to the model can be computed by considering a game between the central server and each client in G_t, as described in the next section.

3. 3. 2. Game Model We suppose the server plays the game independently with each client. The valid actions of the clients are to send good or bad updates while the server can accept or ignore these updates as its valid actions. It should be noted that in our model the players (server and clients) follow a mixed-strategy in which the actions are randomly selected over the set of available actions according to some probability distribution. Afterward, the Nash Equilibrium property is applied to determine the probability of server and client actions. The Nash Equilibrium property is a popular Game Theory concept that describes strategies from which reasonable decision makers should not be deviated to maximize their utility.

The server uses the calculated probability of accepting the updates in our aggregation method as can be seen in Equation (5).

Figure 2 illustrates the normal form of the game including the valid actions and corresponding payoffs for the *i*th player.

$$A_i = \frac{n_i}{N_G}\left|\omega^i{}_{t+1} - \omega_t\right| \tag{6}$$

$$B_i = \frac{n_i}{N_G}\left|\omega^i{}_{t+1}\right| \tag{7}$$

and,

$$x = \begin{cases} 0 & , if\ m^i{}_G \geq m^i{}_B \\ m^i{}_B - m^i{}_G & , if\ m^i{}_G < m^i{}_B \end{cases} \tag{8}$$

where $m^i{}_G$ and $m^i{}_B$ are the number of times that the ith client is in the good and bad set, respectively. Moreover, N_G is the total number of data points in G_t.

As one can see in Figure 2, each client can send good or bad updates while the server can accept or ignore these updates as valid actions. When the *i*th client sends good updates, if the server accepts these updates, it earns a payoff as large as A_i, which indicates the *i*th client contribution to the correction of the previous global model. On the other hand, if the server rejects these good updates, it losses this amount of payoff. Furthermore, we consider the client payoff equal to B_i, that is the client contribution in the global model at the next round, if the server accepts the good updates. It is worth noting that when a client is misdiagnosed several times, the server ignores it forever. Clearly, we should consider the effect of this wrong action of the server in the client payoff when it sends a good model. Therefore, in this situation, we add the term $ln\frac{1}{1+x}$ to the client payoff. In this term, x is related to the number of good model rejection by the server as explained in Equation (8).

Now, if the client is faulty or malicious and sends bad updates, the payoffs of the players can be determined as illustrated in Figure 2. When the server accepts the bad

Server \ Client	Good	Bad
Accept	A_i , B_i	$-A_i$, $-B_i$
Reject	$-A_i + \ln\left(\frac{1}{1+x}\right)$, $-B_i + \ln\left(\frac{1}{1+x}\right)$	0 , 0

Figure 2. The normal form of the game played between the server and the *i*th client

updates, both of them experience negative payoff. On the other hand, if the server rejects these updates, neither side will earn any payoffs.

3. 3. 3. Nash Equilibrium Since in our model, the players follow a mixed-strategy, we can determine the probability of the server and clients actions by applying the Nash Equilibrium property. In other words, there is at least one Nash Equilibrium when we consider mixed-strategy [26]. A mixed strategy Nash Equilibrium involves at least one player playing a randomized strategy and no player is able to increase his or her expected payoff by playing an alternate strategy. At the mixed Nash Equilibrium, both players should be indifferent between their two strategies. Therefore, if the server is using a mixed strategy, it must be indifferent between accepting and rejecting the updates. So, we can write:

$$q_t{}^i A_i - A_i\left(1 - q_t{}^i\right) = q_t{}^i\left(-A_i + ln\left(\tfrac{1}{1+x}\right)\right) \tag{9}$$

and therefore we can calculate the probability of sending good updates by the *i*th client ($\boldsymbol{q_t}^{\boldsymbol{i}}$), when it plays Nash Equilibrium, as follows:

$$\boldsymbol{q_t}^{\boldsymbol{i}} = \frac{A_i}{3A_i - ln\frac{1}{1+x}} \tag{10}$$

On the other hand, when the server plays Nash Equilibrium, the *i*th client should be indifferent between its two actions. So, in a similar way, the probability of accepting the *i*th client updates ($\boldsymbol{p_t}^{\boldsymbol{i}}$) by the server, when it plays Nash Equilibrium, can be derived as follows:

$$\boldsymbol{p_t}^{\boldsymbol{i}} = \frac{B_i - ln\frac{1}{1+x}}{3B_i - ln\frac{1}{1+x}} \tag{11}$$

Finally, the server uses $\boldsymbol{p_t}^{\boldsymbol{i}}$ in Equation (5) when it aggregates the received updates.

It is worth noting that this game has also two pure strategy Nash Equilibriums, i.e. (Good, Accept) and (Bad, Reject), which are quite obvious.

4. EXPERIMENTS

In this section, we report on a detailed experimental study that examines robustness and efficiency of our robust federated learning method. The objective of our experiments is to evaluate the robustness and efficiency of our approach for estimating the global model based on the model received from the clients in the presence of faults.

4. 1. Experimental Environment We conducted experiments on three datasets: CIFAR-10 [27], MNIST [28] and SPAMBASE. CIFAR-10 consists of 60000 32×32 color images in 10 classes while MNIST has 70000 28×28 handwritten digits in 10 classes. For

CIFAR we used VGG-11 which is a familiar convolutional neural network [29] and for MNIST we trained DNNs with $784 \times 512 \times 256 \times 10$ with learning rate 0.1 and Dropout probability 0.5. Also, the hidden layers and output layer activation functions are Leaky ReLU and Softmax, respectively. Furthermore, for SPAMBASE we trained DNNs with $54 \times 100 \times 50 \times 1$ with learning rate 0.05 and Dropout probability 0.5. Also, the hidden layers and output layer activation functions are Leaky ReLU and Sigmoid, respectively. We considered gradient descent as the optimization method where the batch size and epoch number are 200 and 10, respectively. For all the simulations, we set $\alpha = 5$ and we consider the number of clients 10 and 100. Moreover, we assume that all the clients are selected to send updates for the server, i. e. $M_t = N$.

In this paper, we consider four different scenarios, namely, *clean*, *byzantine*, *flipping*, and *noisy* to evaluate our proposed method. For the clean scenario, all of the clients send good updates to the server. In the *byzantine* case, some of the clients are bad and send updates that are significantly different from the updates sent by the good clients. In this case, we consider a Gaussian distribution with mean zero and isotropic covariance matrix with standard deviation 20. For the third scenario that is *flipping*, we set all the labels of data points, used by the selected bad clients to train the model, to zero. Finally, in *noisy* case, we add uniform noise to all the pixels of the noisy clients.

4. 2. Evaluation Results In this section, we compare our proposed method (GFA) with the previous works, namely, Multi-KRUM (MKRUM) [8], Federated Averaging (FA) [1], and COMED [12]. Figures 3, 5, and 7 illustrate the test accuracy of all the algorithms for 10 clients as a function of the number of iterations on the CIFAR-10, MNIST, and SPAMBASE datasets, respectively. In addition, Figures 4, 6, and 8 illustrate the test accuracy of all the algorithms for 100 clients as a function of the number of iterations on the CIFAR-10, MNIST, and SPAMBASE datasets, respectively. According to these figures, we can analyze the convergence of these algorithms. As can be observed, the proposed algorithm converges for all the four scenarios over both datasets while other algorithms do not converge in at least one of the eight cases. For example, the FA and COMED algorithms do not converge for flipping scenarios on the CIFAR-10 while MKRUM and COMED have the same convergence problem on the MNIST. In addition, in the worst case, our algorithm converges after a maximum of 30 iterations.

In addition, Figure 9 illustrates test accuracy as a function of the number of iterations for different values of α on two datasets, namely, CIFAR-10 and SPAMBASE. Accordingly, we set $\alpha = 5$ because it leads to the best accuracy.

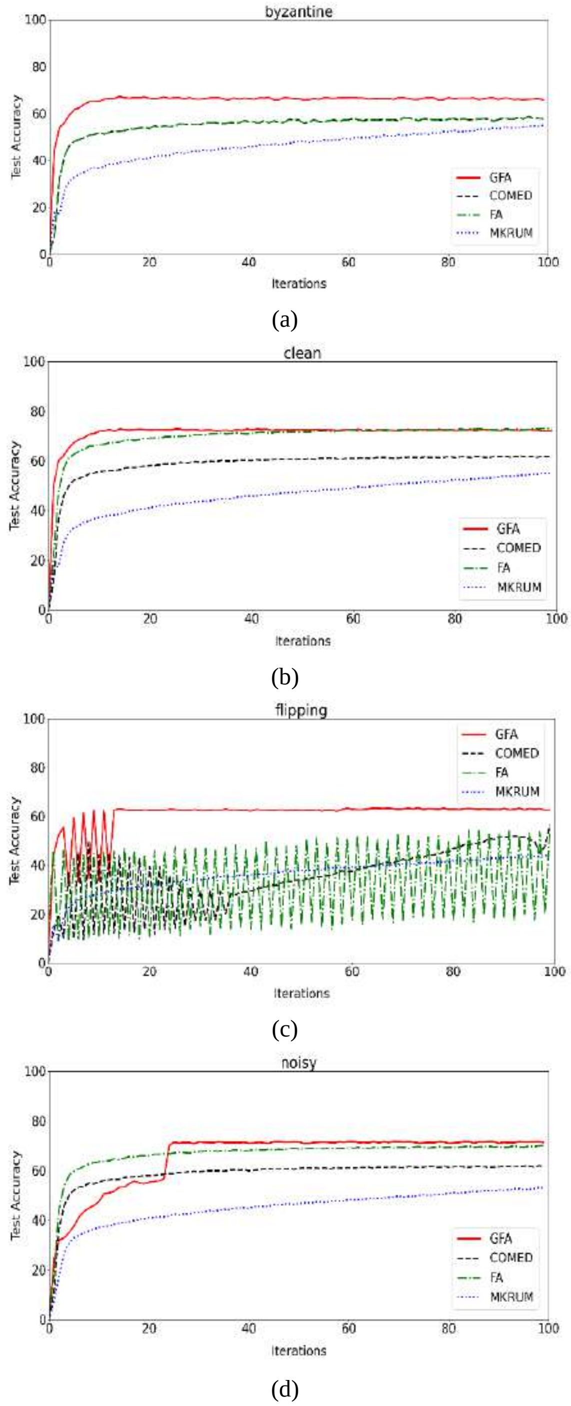
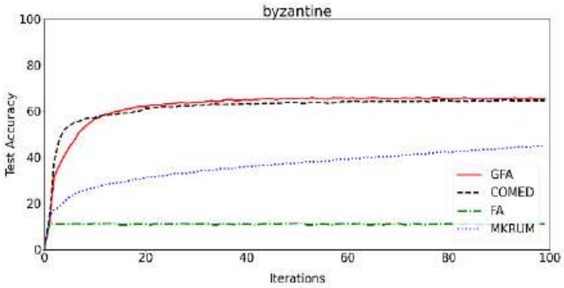
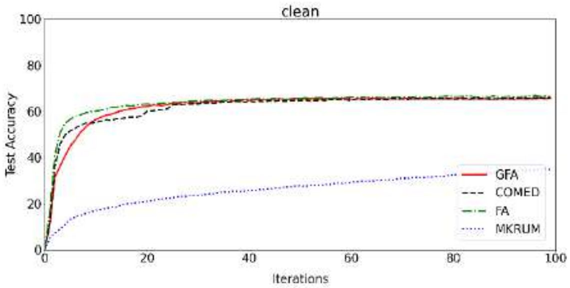


Figure 3. Test accuracy (%) as a function of the number of iterations for 10 clients and for different algorithms on CIFAR-10 for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients

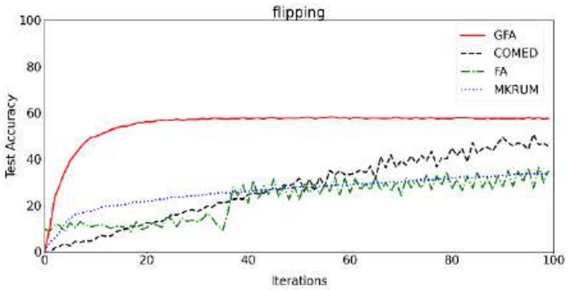
Even in the cases where the other methods converge, the proposed GFA algorithm is ultimately more accurate. Tables 2, 3, and 4 compare the ultimate test accuracy of the different algorithms over CIFAR-10, MNIST, and SPAMBASE, respectively. Accordingly, for CIFAR-10,



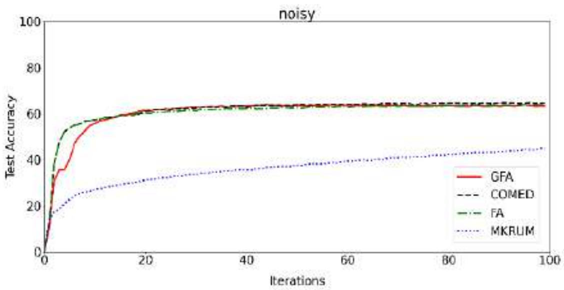
(a)



(b)



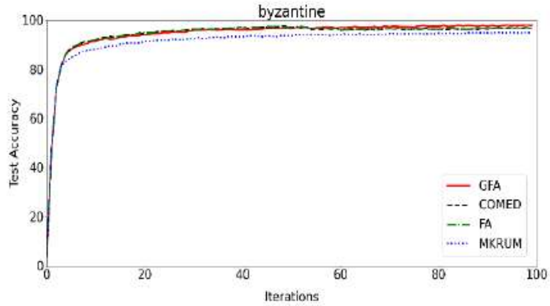
(c)



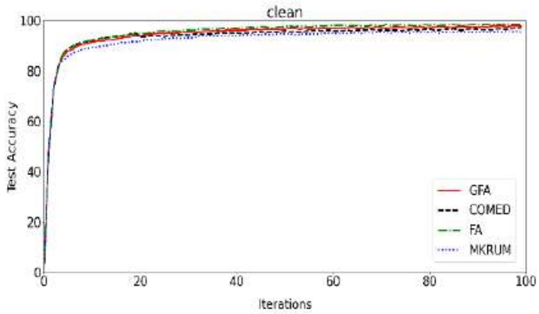
(d)

Figure 4. Test accuracy (%) as a function of the number of iterations for 100 clients and for different algorithms on CIFAR-10 for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients

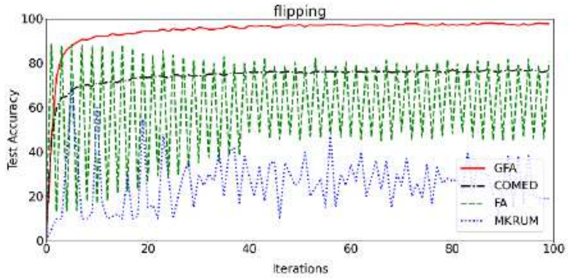
the test accuracy of the proposed algorithm is at least 14%, 15.8%, and 2.3% better than the others for byzantine, flipping, and noisy scenarios, respectively. For this dataset and in the case of clean scenario, the results show that the accuracy of the GFA is only 0.27% less than the standard FA algorithm. On the other hand, for MNIST, the simulations indicate a similar situation where the accuracy of our method is at least 0.4%, 27%, and 2.6% higher than the byzantine, flipping, and noisy scenarios, respectively. Again, the accuracy of the FA algorithm is a little (0.8%) better than the GFA algorithm in the case of clean.



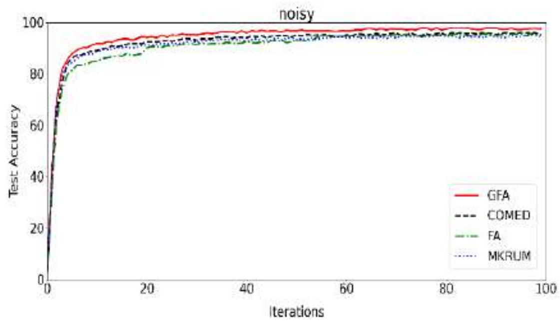
(a)



(b)

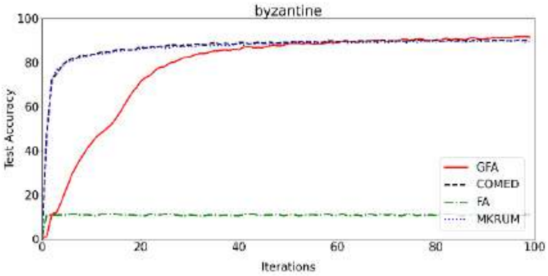


(c)

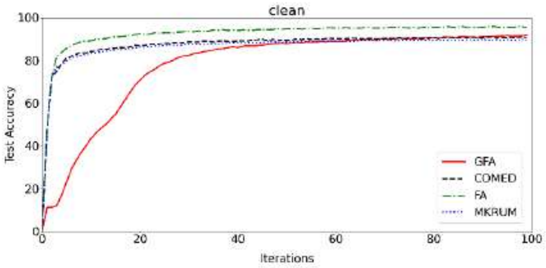


(d)

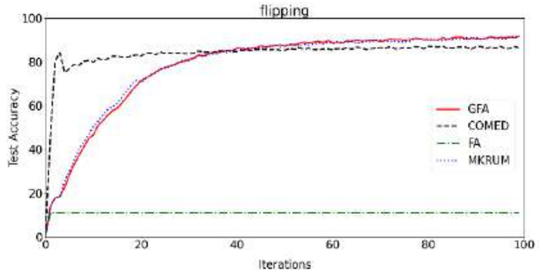
Figure 5. Test accuracy (%) as a function of the number of iterations for 10 clients and for different algorithms on MNIST for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients



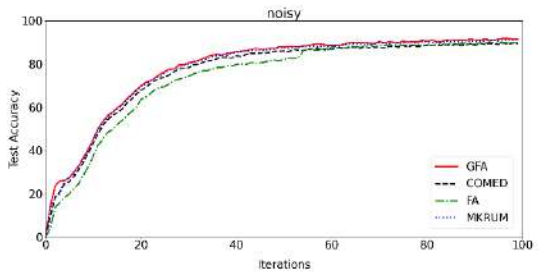
(a)



(b)

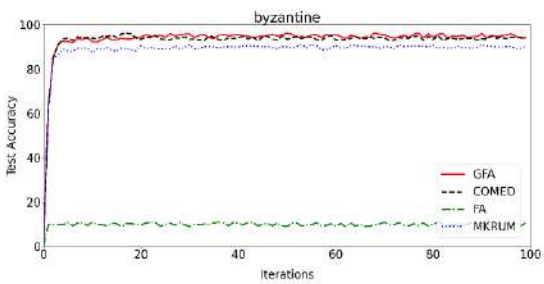


(c)

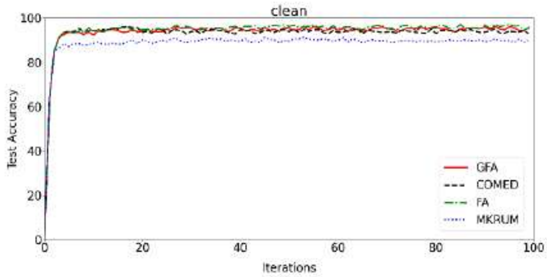


(d)

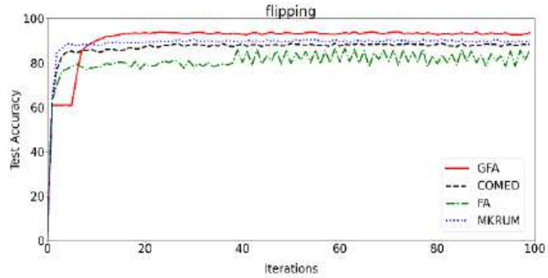
Figure 6. Test accuracy (%) as a function of the number of iterations for 100 clients and for different algorithms on MNIST for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients



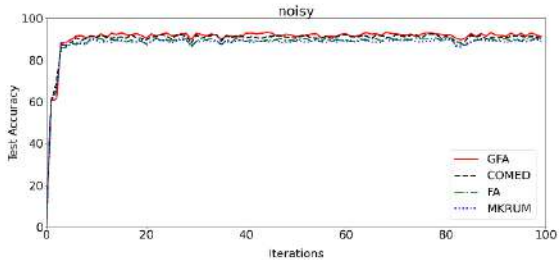
(a)



(b)

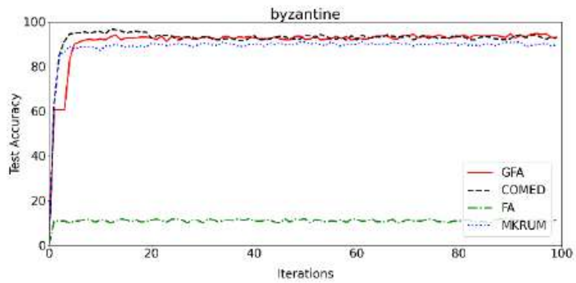


(c)

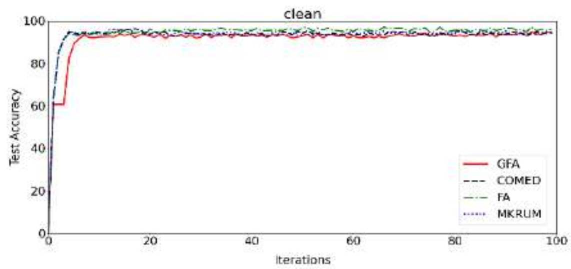


(d)

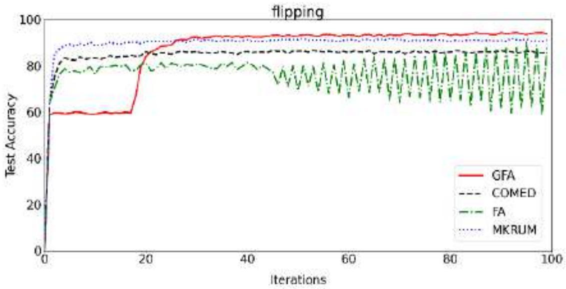
Figure 7. Test accuracy (%) as a function of the number of iterations for 10 clients and for different algorithms on SPAMBASE for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients



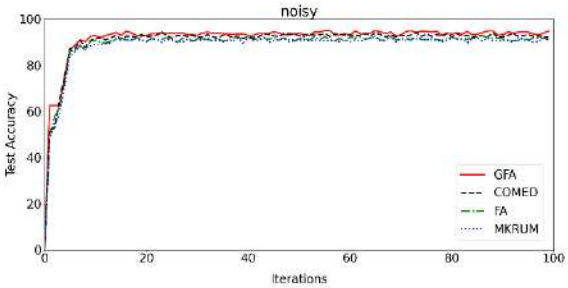
(a)



(b)

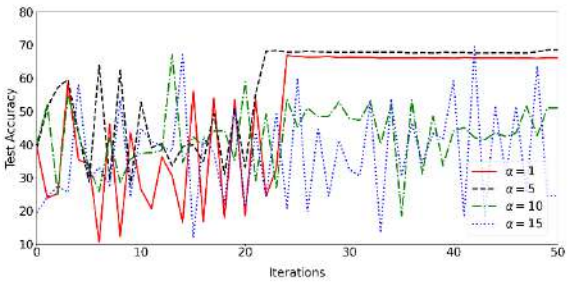


(c)

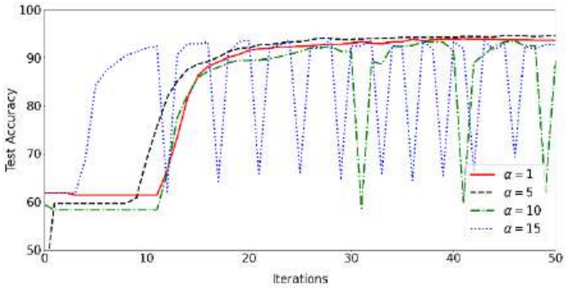


(d)

Figure 8. Test accuracy (%) as a function of the number of iterations for 100 clients and for different algorithms on SPAMBASE for, a) clean (all benign clients), b) byzantine, c) flipping, and d) noisy clients



(a)



(b)

Figure 9. Test accuracy (%) as a function of the number of iterations for different values of α on a) CIFAR-10 and b) SPAMBASE datasets

Furthermore, we investigated the detection rate for bad clients in Table 5. According to this table, for all of the scenarios containing malicious clients, the proposed algorithm in this paper can detect 100% of the bad clients for both datasets.

TABLE 2. The test accuracy of different algorithms for the CIFAR-10 dataset

Algorithm	Clean	Byzantine	Flipping	Noisy
GFA	72.83	67.06	63.48	71.69
FA	73.03	52.49	54.78	70.02
COMED	62.06	58.63	56.84	62.09
MKRUM	54.76	55.32	43.93	53.13

TABLE 3. The test accuracy of different algorithms for the MNIST dataset

Algorithm	Clean	Byzantine	Flipping	Noisy
GFA	98.09	98.01	98.23	98.09
FA	98.89	10.29	77.17	95.57
COMED	97.02	97.43	91.13	95.96
MKRUM	96.01	95.07	70.00	95.22

TABLE 4. The test accuracy of different algorithms for SPAMBASE dataset

Algorithm	Clean	Byzantine	Flipping	Noisy
GFA	96.42	96.65	94.18	93.50
FA	97.19	10.91	86.98	91.13
COMED	95.74	96.24	88.97	92.48
MKRUM	91.64	91.53	90.88	90.58

TABLE 5. The detection rate of GFA algorithm for bad clients on MNIST, CIFAR-10, and SPAMBASE datasets

Dataset	Byzantine	Flipping	Noisy
MNIST	100%	100%	100%
CIFAR10	100%	100%	100%
SPAMBASE	100%	100%	100%

5. CONCLUSION

In this paper, we introduced a game-based robust federated averaging algorithm to detect and discard bad updates provided by the clients. The proposed method uses an adaptive averaging method, in an iteration manner, to highlight the effect of the good updates sent by the majority of the clients. At the end of this iterative algorithm, a trustworthiness is assigned to each client that can be used to put the client in one of the good or bad sets. Finally, the server considers the probability of providing good updates by the clients to the model. These probabilities can be computed by considering a mixed-strategy game between the central server and each client that exists in the good client set. The valid actions of the clients are to send good or bad updates while the server can accept or ignore these updates. By employing the Nash Equilibrium property, the server determines the clients' probability to provide good updates to the model.

In experiments, we considered four scenarios, *clean*, *byzantine*, *flipping*, and *noisy* that were evaluated on MNIST, SIFAR-10, and SPAMBASE datasets for 10 and 100 clients. For all of the scenarios and both datasets, our algorithm converges after a maximum of 30 iterations. It should be noted that in all cases, 100% of the bad clients can be detected for both datasets. In addition, the test accuracy of the proposed algorithm is at least 15.8% and 2.3% better than the others for *flipping* and *noisy* scenarios, respectively.

In future work, we plan to use the Game Theory to detect the backdoor attack where a malicious client can use model replacement to introduce backdoor functionality into the global model.

6. REFERENCES

1. McMahan, B., Moore, E., Ramage, D., Hampson, S. and y Arcas, B.A., "Communication-efficient learning of deep networks from decentralized data", in Artificial Intelligence and Statistics, PMLR. 1273-1282.
2. Konečný, J., McMahan, H.B., Ramage, D. and Richtárik, P., "Federated optimization: Distributed machine learning for on-device intelligence", arXiv preprint arXiv:1610.02527, Vol., No., (2016).
3. Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D. and Shmatikov, V., "How to backdoor federated learning", in International Conference on Artificial Intelligence and Statistics, PMLR., 2938-2948.
4. Yang, Q., Liu, Y., Chen, T. and Tong, Y., "Federated machine learning: Concept and applications", ***ACM Transactions on Intelligent Systems and Technology (TIST)***, Vol. 10, No. 2, (2019), 1-19. Doi: 10.1145/3298981
5. Bhagoji, A.N., Chakraborty, S., Mittal, P. and Calo, S., "Analyzing federated learning through an adversarial lens", in International Conference on Machine Learning, PMLR. 634-643.
6. Li, T., Sahu, A.K., Talwalkar, A. and Smith, V., "Federated learning: Challenges, methods, and future directions", ***IEEE Signal Processing Magazine***, Vol. 37, No. 3, (2020), 50-60. Doi: 10.1109/msp.2020.2975749
7. Sattler, F., Wiedemann, S., Müller, K.-R. and Samek, W., "Robust and communication-efficient federated learning from non-iid data", ***IEEE Transactions on Neural Networks and Learning Systems***, Vol. 31, No. 9, (2019), 3400-3413. Doi: 10.1109/tnnls.2019.2944481
8. Blanchard, P., El Mhamdi, E.M., Guerraoui, R. and Stainer, J., "Machine learning with adversaries: Byzantine tolerant gradient descent", in Proceedings of the 31st International Conference on Neural Information Processing Systems. Vol., No., (Year), 118-128.
9. Damaskinos, G., El Mhamdi, E.M., Guerraoui, R., Guirguis, A.H.A. and Rouault, S.L.A., "Aggregathor: Byzantine machine learning via robust gradient aggregation", in The Conference on Systems and Machine Learning (SysML), 2019. Vol., No. CONF, (Year).
10. Mhamdi, E.M.E., Guerraoui, R. and Rouault, S., "The hidden vulnerability of distributed learning in byzantium", arXiv preprint arXiv:1802.07927, (2018).
11. Nash, J., "Non-cooperative games", Annals of Mathematics, (1951), 286-295.
12. Konečný, J., McMahan, H.B., Yu, F.X., Richtárik, P., Suresh, A.T. and Bacon, D., "Federated learning: Strategies for improving communication efficiency", arXiv preprint arXiv:1610.05492, (2016).
13. McMahan, H.B., Moore, E., Ramage, D. and y Arcas, B.A., "Federated learning of deep networks using model averaging", arXiv preprint arXiv:1602.05629, (2016).
14. Chen, M., Mathews, R., Ouyang, T. and Beaufays, F., "Federated learning of out-of-vocabulary words", arXiv preprint arXiv:1903.10635, (2019).
15. Hard, A., Rao, K., Mathews, R., Ramaswamy, S., Beaufays, F., Augenstein, S., Eichner, H., Kiddon, C. and Ramage, D., "Federated learning for mobile keyboard prediction", arXiv preprint arXiv:1811.03604, (2018).
16. Wang, Y., "Co-op: Cooperative machine learning from mobile devices", , (2017).
17. Yin, D., Chen, Y., Kannan, R. and Bartlett, P., "Byzantine-robust distributed learning: Towards optimal statistical rates", in International Conference on Machine Learning, PMLR., 5650-5659.
18. Xie, C., Koyejo, S. and Gupta, I., "Zeno: Distributed stochastic gradient descent with suspicion-based fault-tolerance", in International Conference on Machine Learning, PMLR., 6893-6901.
19. Sun, Z., Kairouz, P., Suresh, A.T. and McMahan, H.B., "Can you really backdoor federated learning?", arXiv preprint arXiv:1911.07963, (2019).
20. Kang, J., Xiong, Z., Niyato, D., Yu, H., Liang, Y.-C. and Kim, D.I., "Incentive design for efficient federated learning in mobile networks: A contract theory approach", in 2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS), IEEE., 1-5. Doi: 10.1109/vts-apwcs.2019.8851649
21. Feng, S., Niyato, D., Wang, P., Kim, D.I. and Liang, Y.-C., "Joint service pricing and cooperative relay communication for federated learning", in 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), IEEE., 815-820. Doi: 10.1109/ithings/greencom/cpscom/smartdata.2019.00148
22. Zou, Y., Feng, S., Niyato, D., Jiao, Y., Gong, S. and Cheng, W., "Mobile device training strategies in federated learning: An evolutionary game approach", in 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), IEEE., 874-879. Doi: 10.1109/ithings/greencom/cpscom/smartdata.2019.00157
23. Limam, N. and Boutaba, R., "Assessing software service quality and trustworthiness at selection time", ***IEEE Transactions on Software Engineering***, Vol. 36, No. 4, (2010), 559-574. Doi: 10.1109/tse.2010.2
24. Rehman, A.U., Jiang, A., Rehman, A. and Paul, A., "Weighted based trustworthiness ranking in social internet of things by using soft set theory", in 2019 IEEE 5th International Conference on Computer and Communications (ICCC), IEEE., 1644-1648. Doi: 10.1109/iccc47050.2019.9064242

25. De Kerchove, C. and Van Dooren, P., "Iterative filtering in reputation systems", ***SIAM Journal on Matrix Analysis and Applications***, Vol. 31, No. 4, (2010), 1812-1834. Doi: 10.1137/090748196

26. Myerson, R.B., "Game theory, Harvard university press, (2013). Doi: 10.1002/9781118547168

27. Krizhevsky, A. and Hinton, G., "Learning multiple layers of features from tiny images", Vol., No., (2009).

28. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., "Gradient-based learning applied to document recognition", ***Proceedings of the IEEE***, Vol. 86, No. 11, (1998), 2278-2324. Doi: 10.1109/5.726791

29. Liu, S. and Deng, W., "Very deep convolutional neural network based image classification using small training sample size", in 2015 3rd IAPR Asian conference on pattern recognition (ACPR), IEEE., 730-734. Doi: 10.1109/acpr.2015.7486599

Persian Abstract

چکیده

با استفاده ازیادگیری فدراسیونی قابلیت تجمیع مدلهای آموزش یافته بر روی تعداد زیادی از کلاینتها از طریق ارسال این مدلها به سرور مرکزی فراهم می‌گردد. این درحالیست که همچنان حریم خصوصی کلاینتها حفظ خواهد شد، زیرا تنها مدلها به سرور ارسال می‌شود. روشهای یادگیری فدراسیونی به شدت درمعرض حملات قرار دارند. در این مقاله ما یک الگوریتم میانگین‌گیری مقاوم براساس نظریه بازیها ارائه میکنیم. ما فرآیند میانگین‌گیری را با یک بازی با سناریوی میکس که در آن هر کلاینت و سرور به عنوان بازیکن می باشند، مدل میکنیم. اعمال مجاز کلاینتها در بازی شامل ارسال بروزرسانی‌های خوب و بد و نیز اعمال سرور شامل پذیرش یا رد این بروزرسانی‌ها می باشند. نتایج آزمایشها نشان میدهد به کار بردن روش میانگین گیری مبتنی بر نظریه بازیها بسیار مقاومتر از روشهای مشابه در مقابل کلاینتها مخرب می باشد. مطابق این نتایج، روش پیشنهادی حداکثر بعداز 30 تکرار همگرا میشود و قادر است 100 درصد از کلاینتهای مخرب را تشخیص دهد. همچنین روش ما از نظر دقت به ترتیب حداکثر 15.8 درصد و 2.3 درصد بهتر از روشهای پیشین برای دو سناریوی فلیپینگ و نویزی می باشد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 843-853

International Journal of Engineering

Journal Homepage: www.ije.ir

Proposing a New Image Watermarking Method Using Shearlet Transform and Whale Optimization Algorithm

M. Saadati[a], J. Vahidi*[b], V. Seydi[a], P. Sheikholharam Mashhadi[a]

[a] Department of Computer Engineering, Artificial Intelligence, South Tehran Branch, Islamic Azad University, Tehran, Iran
[b] Department of Applied Mathematics, Iran University of Science and Technology, Tehran, Iran

PAPER INFO

Paper history:
Received 14 December 2020
Received in revised form 10 February 2021
Accepted 16 February 2021

Keywords:
Shearlet Transform
Whale Optimization Algorithm
Singular Value Decomposition
Image Watermarking

ABSTRACT

Digital watermarking is a method for data hiding that ensures the security of multimedia data. The watermark can be a digital image or data stored within digital content. The Shearlet transform, a multi-resolution and multi-directional conversion, can be used for watermarking in digital images. Due to its superior features, this conversion can increase the efficiency of applications such as watermarking of images. In this paper, Shearlet and the Singular Value Decomposition (SVD) transforms are used with the Whale optimization algorithm to obtain the most appropriate scaling factor in the watermark extraction step after applying different types of image processing attacks. Shearlet transform has more transparency than traditional converts. The SVD transform also increases the robustness of watermarking operations. The results of different experiments show that the new method presented in this paper, in terms of robustness and imperceptibility compared to the evaluated methods, performs better than most image processing attacks.

doi: *10.5829/ije.2021.34.04a.10*

1. INTRODUCTION

In recent years, computer networks' dramatic growth has facilitated multimedia data dissemination such as images, videos, and audios. Multimedia data can be easily copied and published without the owner's consent. There are different ways to prevent unauthorized disseminations, one of which is digital watermarking. A watermark can be a logo, digital signature, etc., inserted in the host image. Therefore, the image owner can prove the suspicious image from the watermarking image via retrieving the watermark. The main purpose of digital watermarking is embedding the watermark information invisible and usually resistant in digital content [1]. In recent decades many new techniques and concepts have been introduced in the field of image watermarking. They have been used as a tool for detecting changes in digital images as well as image authentication.

A digital watermark is a visible or invisible identification code. It may contain some information about the legal recipient or author of the original data and copyright laws in the form of textual or visual data stored. This digital watermark can be identified or extracted and later used to claim real ownership of the data. The lack of a watermark in the previously watermarking image means the digital data content has changed.

Discrete Cosine Transforms (DCT) [2-4], Discrete Wavelet Transforms (DWT) [5-8], and the Singular Value Decomposition (SVD) Transform [5-14] are the most commonly used frequency domain transformations for watermarking. According to the papers, some hybrid watermarking techniques using the SVD transform can also be mentioned [4-11, 15-17]. Several watermarking studies that use optimization and evolutionary algorithms are Genetic Algorithm (GA) [18-20], Particle Swarm Optimization (PSO) Algorithm [21, 22], and Bee Colony (BC) Algorithm [2, 10, 15, 23] and Wolf Optimization Algorithm [24]. Abdelhakim et al. proposed a resistant watermarking method using the DCT conversion and bee colony algorithm [2]. Seardean et al. proposed a watermarking method in the Wavelet transform domain [24]. Mishra et al. developed a resistant watermarking

*Corresponding Author Institutional Email: jvahidi@iust.ac.ir (J. Vahidi)

Please cite this article as: M. Saadati, J. Vahidi, V. Seydi, P. Sheikholharam Mashhadi, Proposing a New Image Watermarking Method Using Shearlet Transform and Whale Optimization Algorithm , International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 843-853

method in the DWT-SVD transform domain [7]. Ali et al. used the Differential Evolution (DE) algorithm to find the appropriate scaling factor in the DWT-SVD algorithm [5]. Block-based watermarking was suggested by Makbol et al. using Wavelet transform and the SVD transform [8]. Fazli and Moeini developed a resistant watermarking method utilizing a fusion of Wavelet transform and DCT, and the SVD transform [17]. Zhao et al. proposed a new approach using Shearlet transform and the entropy rate [25]. Xiang-yang et al. proposed a watermarking method that employs watermarking in the sub-band with the highest amount of energy [26].

The rest of the paper is organized as follows. In Section 2, we review the Materials and Methods. Section 3 presents the proposed method in this paper. Section 4 illustrates the evaluation of results. Section 5 shows the result of simulations and experiments. Finally, Section 6 concludes this paper.

2. MATERIALS AND METHODS

2. 1. Classification of Digital Watermarking Methods Watermark embedding techniques can be divided into two main groups. Below is a brief description of the features of these groups.

In spatial domain techniques, the watermark is added directly to the original image by changing its pixel values. These algorithms are simple, fast, and have a high embedded capacity [27]. These techniques are susceptible to changes that can damage the watermark [28, 29]. Spatial-based techniques' disadvantages are more than their advantages that cannot withstand many attacks, such as added noise and compression methods [30]. Techniques for this class include Least Significant Bit (LSB) [31], Local Binary Pattern (LBP) [32], and Histogram modification [33].

Compared to spatial domain methods, in transform domain techniques, the goal is to incorporate watermarking into image coefficients. The host image is first transmitted from the spatial domain to the transform domain with a reversible conversion and after inserting the tokens with a suitable embedding algorithm. An important issue in this area is to choose the best place for watermarking to avoid image distortion. The transform field is resistant to change, attack and the watermark can be invisibly embedded in the image [28, 29]. Techniques of this method include DCT [34], Discrete Fourier Transform (DFT) [35], and DWT [6].

2. 2. Shearlet Transform A useful feature of Wavelets is their ability to find the image features in horizontal, vertical, and diagonal directions. However, Wavelets are not capable of identifying image-oriented features in other directions because of isotropic bases. To solve this problem and find features in different directions, we can use other multi-scale directional transformations. A qualified feature of these transformations in watermarking operations is that we can embed the watermark in differently oriented sub-bands of image. These multi-scale and multi-resolution transformations include Wavelet, Contourlet, and Shearlet transform. The Contourlets have bases made of a combination of directional and multi-scale filter banks [36].

In 2005, a new approach was proposed by Labate et al. [37], called Shearlet, which obtained a near-optimal approximation [38]. The Shearlet display has been introduced as a multi-directional Wavelet method that provides almost optimal features in recent years. This new representation is based on a resistant and simple mathematical framework that provides a flexible tool for geometric representation of multi-dimensional data and is also more natural for implementation. The Shearlet method is also related to multi-resolution analysis [36]. Figure 1 shows the frequency domain of the Shearlet transform [39].

The continuous 2D Shearlet transform (Shearlet's mother) in $\psi \in L^2(R^2)$ space is defined as Equation (1):

$$\begin{aligned}\psi_{a,s,t}(x) &= a^{-3/4}\psi(A_a^{-1}S_s^{-1}(x-t)) \\ &= a^{-3/4}\psi\left(\begin{pmatrix}\frac{1}{a} & \frac{-s}{a} \\ 0 & \frac{1}{\sqrt{a}}\end{pmatrix}(x-t)\right), a>R^+, s\in R, t\in R^2\}\end{aligned} \quad (1)$$

where ψ is called the generator function, A and S are 2x2 invertible matrices. A is the parabolic scaling matrices, and S is the Shear matrices [37], as:

$$A_a = \begin{bmatrix} a & 0 \\ 0 & a^{1/2}\end{bmatrix} \quad , \quad S_s = \begin{bmatrix} 1 & s \\ 0 & 1\end{bmatrix} \quad (2)$$

Applying the Fourier transform to the continuous Shearlet transform is obtained using Equation (3):

$$SH_{\psi}(f) = \langle f, \psi_{a,s,t} \rangle \quad (3)$$

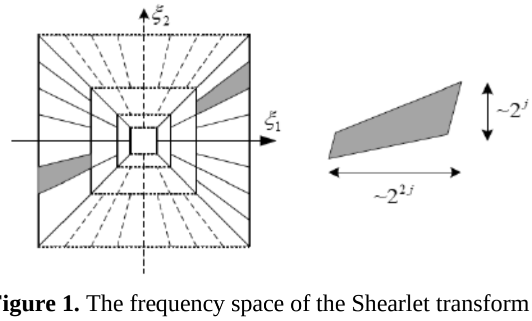


Figure 1. The frequency space of the Shearlet transform [39]

The digital image is considered as sampled functions on the network with

$$\{(\frac{m_1}{M}, \frac{m_2}{N}) : (m_1, m_2) \in I\}$$

$I = \{(m_1, m_2) : m_1 = \circ, ..., M-1, m_2 = \circ, ..., N-1\}$.

We have acquired a discrete Shearlet by discretizing the parameters *a*, *s*, and *t*. Suppose $j = \left\lfloor \frac{1}{2} \log_2 N \right\rfloor$ is the number of levels taken into account. A discrete Shearlet transform must be obtained. The scale, shear, and transfer parameters are discretized :

$$a_j = 2^{-2j} = \frac{1}{4^j} \qquad j = \circ, ..., j_\circ - 1$$

$$s_{j,k} = k 2^{-j} \qquad 2^{-j} \le k \le 2^j$$

$$t_m = (\frac{m_1}{M}, \frac{m_2}{N}) \qquad m \in I$$

With these symbols, discrete Shearlets are defined as Equation (4):

$$\psi_{j,k,m}(x) = \psi(A_{a_j}^{-1} S_{s_{j,k}}^{-1} (x - t_m)) \tag{4}$$

The discrete Shearlet transform is defined as Equation (5):

$$SH_\psi(f) = \left\langle f, \psi_{j,k,m}(x) \right\rangle \tag{5}$$

In this paper, a NonSubsampled Shearlet Transform (NSST) is used. The size of all sub-bands of the image is equal, and its proper feature is being shifted invariant.

2. 3. Singular Value Decomposition A new conversion watermarking method called SVD has been studied in recent years. In the SVD transform, one matrix can be divided into three matrices that are the same size as the original matrix. If *A* is a square image, its corresponding matrix is represented as $A \in R^{n \times n}$, where *R* represents the range of real numbers, so the SVD of matrix *A* is as Equation (6) [40]:

$$A = USV^T = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,N} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ u_{N,1} & u_{N,2} & \cdots & u_{N,N} \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \lambda_N \end{bmatrix} \times \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,N} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,N} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ v_{N,1} & v_{N,2} & \cdots & v_{N,N} \end{bmatrix}^T \tag{6}$$

In the above equation $U \in R^{n \times n}$ and each column constitutes the eigenvectors of the matrix AA^T. These special vectors are called left eigenvectors. Also $V \in R^{n \times n}$ represents the matrix that each column represents the eigenvectors of the matrix AA^T. These eigenvectors are called right eigenvectors, and V^T represents the matrix transpose *V*, which is an identity matrix *n x n*. $S \in R^{n \times n}$ is a diagonal matrix with real nonnegative layers on the original diameter. Each element has a single non-descending matrix *A* value. So all of its elements are zero except for the original diameter. λ is a singular value and is defined as follows:

$$\lambda_1 \ge \lambda_2 \ge ... \lambda_r \ge \lambda_{r+1} = ... = \lambda_N = 0$$

The SVD is a desirable method for matrix decomposition in which the maximum signal energy is packaged using several coefficients. The use of the SVD in digital image watermarking has many advantages, some of which are as follows [41]:

1. Maximum image energy is obtained by the largest singular value coefficients of the image.
2. Singular image values are very stable. When small damage to the image occurs, its values do not change much, so it is highly resistant to various attacks.
3. Image transparency remains almost constant even after embedding the watermark in singular image values.

2. 4. Whale Optimization Algorithm The scaling factor plays an essential role in watermarking operation. An enormous scale factor value makes the watermark image more robust to attacks. A small scale factor value also makes the watermark image more imperceptible. You have to be smart to choose the right scale factor and balance the watermarking image's resistance and transparency. In this paper, we use Whale Optimization Algorithm (WOA) to select the best scale factor.

The WOA is one of the nature-inspired metaheuristic optimization algorithms [42]. It has shown good performance due to its high convergence speed.

Humpback whales can detect hunting location and circle around them since the optimal position is unknown already in search space. As a result, the WOA algorithm assumes that the current candidate's best answer is the optimal target or the solution is close to it. In the following, other search agents try to update their positions according to the best search factor, such as Equations (7) and (8).

$$\vec{D} = \left| \vec{C}.\vec{X}^*(t) - \vec{X}(t) \right| \tag{7}$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A}.\vec{D} \tag{8}$$

where t represents the current iteration, $\vec{A}$ and $\vec{C}$ represent the coefficient vectors. X^* indicates the position vector of the best solution to date. $\vec{X}$ represents the current location vector. The $\vec{A}$ and $\vec{C}$ vectors are calculated using Equations (9) and (10):

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \tag{9}$$

$$\vec{C} = 2.\vec{r} \tag{10}$$

where $\vec{a}$ linearly decreases from 2 to zero during the iterations, and $\vec{r}$ represents a random vector in the interval [0,1].

Humpback whales operate at the exploitation phase using two Shrinking encircling mechanisms and spiral updating position. The contraction loop mechanism is performed by reducing the value of $\vec{a}$ in Equation (9). The distance between the whale at position (X, Y) and the hunting at position (X*, Y*) is first calculated in the spiral position update phase. In the following, a spiral equation is created between the whale position and the hunting position to simulate the spiral motion of humpback whales with Equation (11):

$$\vec{X}(t+1)=\vec{D}'.e^{bl}.\cos(2\pi l)+\vec{X}^*(t) \tag{11}$$

where $\vec{D}'=|\vec{X}^*(t)-\vec{X}(t)|$ and indicates the distance of the ith whale to the prey (best solution obtained so far), b demonstrates a constant that specifies the shape of the logarithmic spiral, l shows a random number in the range [-1,1].

Humpbacks move around the hunt simultaneously, both in the shape of the shrinking circle and along a spiral-shaped path. We assume that a 50% probability can be used by a contraction loop mechanism or a spiral model to model this synchronous behavior and update the whale's position during optimization. The mathematical model is shown in Equation (12):

$$\vec{X}(t+1)=\begin{cases}\vec{X}^*(t)-\vec{A}.\vec{D} & if\ \ p<0.5\\ \vec{D}'e^{bl}.\cos(2\pi l)+\vec{X}^*(t) & if\ \ p\geq 0.5\end{cases} \tag{12}$$

where p represents a random number in the interval [0,1], vector $\vec{A}$ change-based approach also can be used for hunting search. Humpback whales perform a random search based on each other's position. Therefore, we use random values of $\vec{A}$ greater than one or smaller than minus one to force the search agent to avoid a reference whale. Unlike the extraction phase, we update a search agent's position in the exploration phase based on the randomly selected search agent. This mechanism and $|\vec{A}|>1$ emphasize exploration and allow the WOA algorithm to perform a global search. The mathematical model of this work is given by Equations (13) and (14):

$$\vec{D}=\left|\vec{C}.\vec{X}_{rand}-\vec{X}\right| \tag{13}$$

$$\vec{X}(t+1)=\left|\vec{X}_{rand}-\vec{A}\vec{D}\right| \tag{14}$$

where $\vec{X}_{rand}$ is the random location vector selected from the current population.

3. PROPOSED METHOD

This paper proposes an invisible watermarking method based on the SVD technique and WOA algorithm in the Non-subsample Shearlet Transform domain. In the suggested algorithm, the host image, the watermark image, and the watermarking image are defined as I, W, and I_w, respectively.

The watermark embedding algorithm in the host image is as follows:

1. Host image I is parsed into three I_R, I_G, and I_B matrices.
2. In matrix I_i ; i=R, G, B, Shearlet transform is executed up to three levels and decomposed into a low-frequency sub-band and some high-frequency sub-bands.
3. In the low-frequency sub-band obtained from the image in step 2, the SVD transform is performed.
4. The dimension of the W watermark image is resized to the sub-dimension of the host image.
5. The watermark image W is split into three W_R, W_G, and W_B matrices.
6. The watermark image is multiplied by the T scale factor. It is aggregated by the matrix of single values (S) obtained from the host image's low-frequency sub-band.
7. The inverse of the SVD transform is obtained by the matrices U and V obtained in step 3 and the matrix S obtained by step 6.
8. The inverse Shearlet transform is executed on the new low-frequency sub-band matrix obtained from step 7. The high-frequency sub-bands obtained from step 2 and the watermarking image I_i ; i=R, G, B is obtained.
9. The three I_R, I_G, I_B matrices obtained from step 8 are merged, and the I_w watermarking is created.
10. The S matrix obtained from level 6 is then stored as a key to retrieving the watermark image.

The watermark extraction algorithm is as follows:

1. Watermarking I_w image is divided into three matrices I_{WR}, I_{WG}, and I_{WB}.
2. In matrix I_W i ; i=R, G, B, Shearlet transform is executed for up to 3 levels and decomposed into a low frequency and some high-frequency sub-bands.
3. In the low-frequency sub-band obtained from step 2, the SVD transform is performed.
4. The matrix S obtained during the watermark's embedding phase (the key to retrieving the watermark image) is called.
5. The inverse of the SVD transform is executed on the U and V matrices obtained from step 3 and the S matrix (watermarking operation key).
6. The inverse Shearlet transform is executed on the low-frequency sub-band matrix obtained from the inverse SVD transform operation obtained from step 5. The rest of the high-frequency sub-bands obtained from step 2 and the watermarking image I_i ; i=R, G, B is obtained.
7. The three I_R, I_G, and I_B matrices obtained from step 6 are merged to create the extracted host image.
8. The S matrix called in step 4 is subtracted from the S matrix obtained in step 3 and then divided into the T scaling factor (alpha), which the WOA algorithm gets its optimal value. (The evaluation function of the WOA algorithm is the Peak Signal to Noise Ratio (PSNR) value between the host and the extracted watermark).

9. The three W_R, W_G, and W_B matrices are merged to create the extracted watermark.

The watermark embedding and extraction algorithm in the host image are shown in Figures 2 and 3, respectively.

4. EVALUATION OF RESULTS

All implementations and results are obtained using MATLAB 2015 software. Two-color images, Airplane and Lena, are used as host images and Peugeot as the watermark image. The used images were afforded from the USC-SIPI database and are standard images for various image processing operations. Figures 4 and 5 show the host images and the watermark image, respectively. The host image size is 128*128 pixels. The watermark is also a 32x32 pixel color image.

In digital watermarking, it is necessary to maintain quality of the original image. Therefore, two criteria of imperceptibility and robustness are needed to evaluate the watermarking and watermark image. These criteria measure the amount of transparency between two images [43]. PSNR is used to measure imperceptibility. The mean square error (MSE) between the two images is measured using Equation (15) [44].

$$MSE(I,I_w)=\frac{1}{MN}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(I_1(i,j)-I_2(i,j))^2 \tag{15}$$

M and N are the dimensions of the image, I_1 and I_2 are the primary and the extracted image, respectively.

PSNR is used to estimate the quality of two images. The high PSNR indicates that the two images are very similar [45]. This criterion is expressed in decibel unit (db.) using Equation (16):

$$PSNR(I_1,I_2)=10*\log 10\frac{MAX_I^2}{MSE} \tag{16}$$

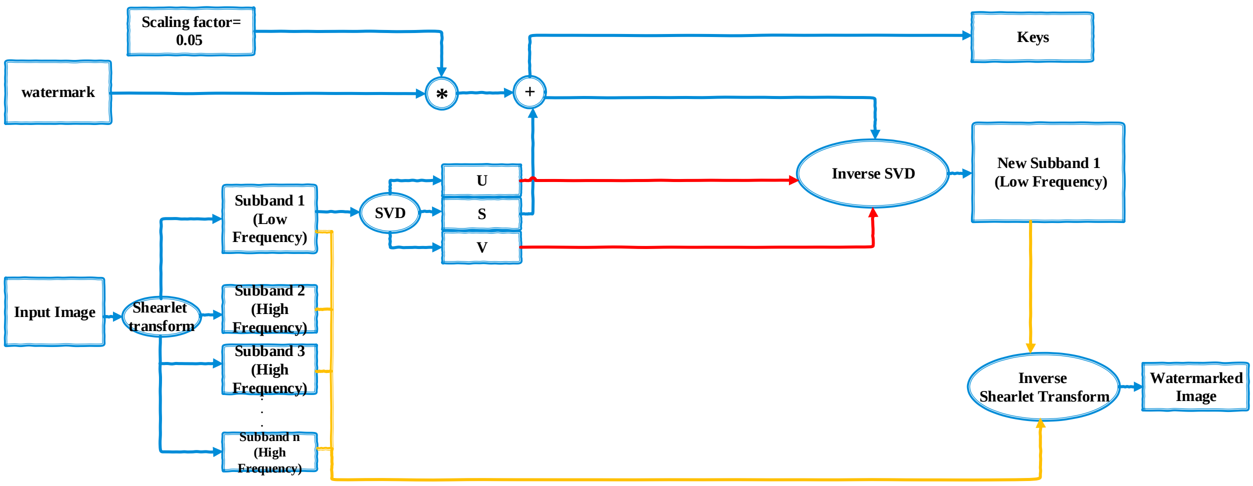


Figure 2. Watermarking embedding algorithm

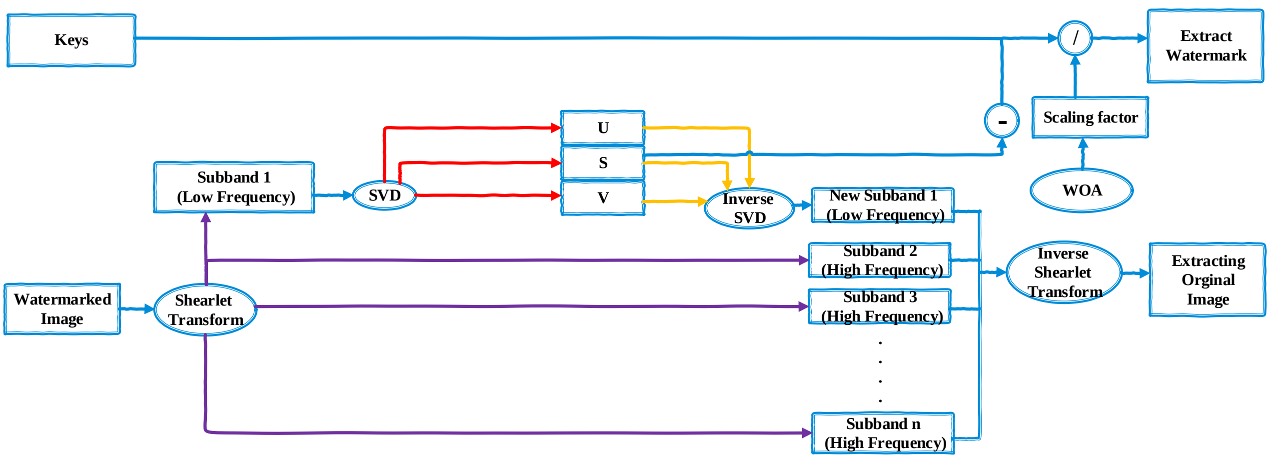


Figure 3. Watermarking extraction algorithm

Figure 4. Images used for the host

Figure 5. Image used for watermarking

MAX_I represents the peak signal in the input image, which is 255 in 8-bit images.

Correlation Coefficient (CRC) can be used to measure the robustness of the extracted watermark. The CRC is used to measure the correlation between the two images. The CRC value ranges from zero to one, calculated using Equation (17) [46].

$$CRC = \frac{\sum_i \sum_j I_1(i,j) I_2(i,j)}{\sqrt{\sum_i \sum_j I_1(i,j)^2 * \sum_i \sum_j I_2(i,j)^2}} \tag{17}$$

I_1 and I_2 are the initial and the extracted image, respectively.

5. EXPERIMENTAL AND COMPARATIVE RESULTS

In this section, results of the proposed method are presented in terms of transparency and robustness according to the experiments. In this method, we have applied Shearlet transform to three levels on the input image. The initial scaling factor value for the watermarking operation was set as 0.05, which is obtained experimentally with relatively good strength and transparency. The WOA algorithm is iterated 100 times.

The proposed method in this paper is compared with similar watermarking techniques using NonSubampled Wavelet Transform (NSWT), NonSubsampled Contourlet Transform (NSCT), DWT_SVD [47], and DST-BSVD [16]. The investigated attacks are as follows: average filter (AF), Gaussian low-pass filter (GP), median filter (MF), Gaussian noise (GN), speckle noise (SN), pepper and salt noise (SP), blurring (BL), motion blur (MB), sharpening (SH), JPEG compression (JPEG), crop (CR), rotation (RO), transition (TR), Vertical flip (FLV), Horizontal flip (FLH).

Comparative results of the extracted watermark transparency based on PSNR criterion and the extracted watermark robustness based on CRC criterion are presented in Table 1. The results of the experiments show that the transparency parameter of the proposed PSNR method is improved on all images compared to the tested methods after AF 3x3, AF 5x5, GP 3x3, GP 5x5, MF 3x3, MF 5x5, RO (5), RO (45), RO (110), FLH, FLV, BL(0.3), BL(0.5), BL(1), MB(15, 45), SH(0.3), SH(0.5), SH(1), JPEG(5), JPEG(20), JPEG(80), JPEG(90), GN(0,0.001), GN(0,0.1), GN(0,0.3), SN(0.001), SN(0.3) and SP(0.3) attacks.

The test results also show that the CRC robustness parameter of the proposed method improved on all images compared to the tested methods after AF 3x3, AF 5x5, GP 3x3 , GP 5x5 , MF 3x3 , MF 5x5 , RO(5) , FLH , FLV , BL(0.3) , BL(0.5) , BL(1) , MB(15,45) , SH(0.3) , SH(0.5) , SH(1) , JPEG(5) , JPEG(20) , JPEG(80) , JPEG(90), GN(0,0.001), GN(0,0.1), GN(0,0.3), SN(0.001) , SN(0.3), and SP(0.001) attacks.

For instance, Lena's watermarking image and the extracted watermark can be seen after applying different image processing attacks in Figure 6.

TABLE 1. Comparison of transparency and robustness of the proposed method with the traditional one

		Airplane Image						Lena Image					
Attacks		NSWT	NSCT	DWT SVD	DST BSVD	NSST WOA	Alpha	NSWT	NSCT	DWT SVD	DST BSVD	NSST WOA	Alpha
No Attack	PSNR	27.5532	27.5422	21.9887	27.6045	27.6082	0.0495	27.5642	27.5374	22.1424	27.6045	27.6722	0.0495
	CRC	0.9976	0.9976	0.9912	0.9976	0.9976		0.9976	0.9976	0.9915	0.9976	0.9976	
AF 3x3	PSNR	-0.0193	20.7062	5.0495	24.4168	24.4186	0.0499	1.3383	21.1577	6.1167	25.9312	25.9561	0.0497
	CRC	0.5612	0.9882	0.7822	0.9949	0.9949		0.617	0.9893	0.8162	0.9964	0.9964	
AF 5x5	PSNR	-6.1806	14.0629	-0.188	20.7052	20.728	0.0506	-5.105	14.5596	0.7797	23.4313	23.4325	0.05
	CRC	0.3434	0.9493	0.6046	0.9881	0.9881		0.3774	0.9544	0.6436	0.9936	0.9936	
GP 3x3	PSNR	9.0892	26.2659	14.7156	27.2263	27.2838	0.0496	10.4572	26.4244	15.9669	27.4432	27.5103	0.0496
	CRC	0.871	0.9967	0.9595	0.9974	0.9974		0.9002	0.9968	0.9691	0.9975	0.9975	

GP 5x5	PSNR	9.034	26.2477	14.6593	27.2223	27.2791	0.0496	10.3999	26.4077	15.9094	27.4415	27.5081	0.0496
	CRC	0.8697	0.9967	0.959	0.9974	0.9974		0.8991	0.9968	0.9687	0.9975	0.9975	
MF 3x3	PSNR	6.854	24.4239	11.3875	27.2733	27.3409	0.0495	8.5258	26.4693	12.1789	27.4907	27.549	0.0495
	CRC	0.8084	0.9949	0.9234	0.9974	0.9974		0.8602	0.9969	0.9353	0.9975	0.9975	
MF 5x5	PSNR	0.927	21.8984	5.5602	26.6743	26.705	0.0496	1.3096	21.9304	5.8633	27.2448	27.2853	0.0495
	CRC	0.5908	0.9909	0.7991	0.997	0.997		0.6147	0.991	0.8083	0.9974	0.9974	
CR (50,50)	PSNR	-0.6669	16.3296	-0.2934	21.0186	21.0333	0.0504	4.6409	21.3338	4.28	17.0051	17.1365	0.0521
	CRC	0.4851	0.968	0.493	0.9889	0.9889		0.7045	0.9897	0.6811	0.9725	0.9725	
RO (5)	PSNR	-1.4583	13.7274	2.6289	15.6693	15.909	0.0533	5.6364	20.0844	8.1568	24.203	24.2561	0.0499
	CRC	0.4235	0.9428	0.6073	0.9634	0.9634		0.7479	0.9862	0.8333	0.9947	0.9947	
RO (45)	PSNR	-9.561	7.6148	-4.1504	16.1511	16.3118	0.0528	-3.7542	13.6737	0.1081	16.1796	16.3986	0.0527
	CRC	0.1379	0.8072	0.2731	0.9669	0.9667		0.343	0.9427	0.5016	0.9671	0.9674	
RO (110)	PSNR	-10.7669	6.2575	-5.1083	10.8244	11.6408	0.0614	-3.8412	13.0419	-0.0082	20.7888	20.8694	0.0505
	CRC	0.1191	0.7586	0.2509	0.8997	0.8989		0.336	0.9343	0.4897	0.9883	0.9885	
TR (5,10)	PSNR	-8.7343	9.8694	-0.3362	5.7629	8.1631	0.0872	-5.6842	12.5935	3.1051	9.187	10.4037	0.0666
	CRC	0.2305	0.88	0.5076	0.7566	0.7565		0.3181	0.9303	0.665	0.8629	0.8629	
TR (10,10)	PSNR	-11.1896	7.3059	-2.7841	3.3138	6.948	0.1153	-8.6026	9.6226	0.2918	6.4978	8.5854	0.0812
	CRC	0.1802	0.8105	0.4145	0.659	0.6589		0.2399	0.875	0.5497	0.7822	0.7823	
TR (10,15)	PSNR	-13.0371	5.364	-4.8295	1.7165	6.3338	0.1442	-10.2198	7.9887	-1.6355	4.9032	7.6904	0.0952
	CRC	0.1513	0.7431	0.348	0.5903	0.5901		0.2063	0.8324	0.4751	0.7231	0.7232	
FLH	PSNR	27.5532	27.5422	21.9887	27.6045	27.68	0.0495	27.5642	27.5374	22.1424	27.6045	27.677	0.0495
	CRC	0.9976	0.9976	0.9912	0.9976	0.9976		0.9976	0.9976	0.9915	0.9976	0.9976	
FLV	PSNR	27.5532	27.5422	21.9887	27.6045	27.6799	0.0495	27.5642	27.5374	22.1424	27.6045	27.68	0.0495
	CRC	0.9976	0.9976	0.9912	0.9976	0.9976		0.9976	0.9976	0.9915	0.9976	0.9976	
BL (0.3)	PSNR	27.2721	27.5422	22.6795	27.6085	27.684	0.0495	27.384	27.5375	22.7695	27.6066	27.6142	0.0495
	CRC	0.9974	0.9976	0.9925	0.9976	0.9976		0.9975	0.9976	0.9927	0.9976	0.9976	
BL (0.5)	PSNR	9.034	26.2477	14.6593	27.2223	27.2796	0.0496	10.3999	26.4077	15.9094	27.4415	27.5043	0.0496
	CRC	0.8697	0.9967	0.959	0.9974	0.9974		0.8991	0.9968	0.9687	0.9975	0.9975	
BL (1)	PSNR	-1.5598	19.1654	3.7335	24.0507	24.0507	0.0499	-0.3246	19.6164	4.795	25.74	25.7613	0.0497
	CRC	0.5009	0.9833	0.7407	0.9945	0.9945		0.5504	0.9849	0.7779	0.9963	0.9963	
MB (15,45)	PSNR	-10.3849	8.3785	-2.8112	13.9633	14.3528	0.0552	-7.7979	10.6649	-0.3996	18.3292	18.386	0.0515
	CRC	0.2275	0.8457	0.4851	0.9472	0.9472		0.2866	0.899	0.5726	0.9796	0.9796	
SH (0.3)	PSNR	27.2895	27.5427	21.548	27.6041	27.6795	0.0495	27.3484	27.5383	21.8401	27.6035	27.6666	0.0495
	CRC	0.9974	0.9976	0.9902	0.9976	0.9976		0.9975	0.9976	0.9909	0.9976	0.9976	
SH (0.5)	PSNR	12.7994	26.7514	12.9629	27.5937	27.6692	0.0495	13.8829	26.729	13.9077	27.5467	27.621	0.0495
	CRC	0.9282	0.9971	0.9284	0.9976	0.9976		0.9431	0.9971	0.942	0.9976	0.9976	
SH (1)	PSNR	2.2872	22.4063	5.302	27.5768	27.6522	0.0495	3.3514	22.6085	6.2613	27.2079	27.2664	0.0496
	CRC	0.5622	0.9919	0.6786	0.9976	0.9976		0.613	0.9923	0.7222	0.9974	0.9974	
JPEG (5)	PSNR	0.1168	17.1102	7.6878	24.2946	24.2964	0.0498	0.6841	18.9329	6.3203	23.9304	23.9319	0.0499
	CRC	0.5055	0.9728	0.8157	0.9948	0.9948		0.5386	0.982	0.7749	0.9943	0.9943	
JPEG (20)	PSNR	5.2668	21.2887	12.0118	27.1215	27.1803	0.0496	7.0401	22.4416	13.4874	27.2512	27.3047	0.0496
	CRC	0.7219	0.9895	0.9151	0.9973	0.9973		0.791	0.992	0.9386	0.9974	0.9974	

JPEG (80)	PSNR	11.4999	25.4888	16.0991	27.5979	27.6697	0.0495	13.3032	26.2067	18.2342	27.608	27.6824	0.0495
	CRC	0.9079	0.9961	0.9652	0.9976	0.9976		0.9372	0.9967	0.9787	0.9976	0.9976	
JPEG (90)	PSNR	14.0842	26.4018	17.9799	27.6153	27.6431	0.0495	15.7238	26.8385	19.7938	27.6108	27.6355	0.0495
	CRC	0.9466	0.9968	0.9774	0.9976	0.9976		0.963	0.9971	0.9853	0.9976	0.9976	
GN (0,0.001)	PSNR	21.6183	27.3805	19.4741	27.6103	27.6748	0.0495	22.6616	27.4792	19.5994	27.5907	27.6624	0.0495
	CRC	0.9904	0.9975	0.9842	0.9976	0.9976		0.9924	0.9975	0.9846	0.9976	0.9976	
GN (0,0.1)	PSNR	-9.893	9.4688	-4.5191	11.4189	12.4969	0.0589	-8.1918	12.9414	-4.0368	15.7908	16.478	0.0526
	CRC	0.1248	0.8672	0.1813	0.9109	0.9181		0.1358	0.9328	0.1761	0.9641	0.9685	
GN (0,0.3)	PSNR	-15.8199	3.3436	-9.2512	4.6152	7.711	0.0952	-12.5428	7.0673	-7.9059	9.3093	11.3599	0.062
	CRC	0.0508	0.6519	0.0654	0.7127	0.7248		0.0607	0.794	0.0576	0.8646	0.8917	
SN (0.001)	PSNR	23.9428	27.4682	20.3124	27.5969	27.6488	0.0495	25.2495	27.5064	21.335	27.5965	27.6447	0.0495
	CRC	0.9944	0.9975	0.987	0.9976	0.9976		0.9959	0.9976	0.9898	0.9976	0.9976	
SN (0.1)	PSNR	-6.8981	12.5193	-1.6489	14.8625	15.2098	0.0541	-3.066	15.8474	1.5968	19.1179	18.6345	0.0513
	CRC	0.2014	0.9274	0.3052	0.9564	0.9569		0.3275	0.9644	0.4785	0.9829	0.9807	
SN (0.3)	PSNR	-14.2862	4.1495	-8.331	5.0541	7.9473	0.0907	-10.0631	8.3206	-5.0202	10.5921	11.4528	0.0618
	CRC	0.0709	0.6865	0.0951	0.73	0.7419		0.1249	0.8358	0.1742	0.8944	0.8944	
SP (0.001)	PSNR	26.4202	27.5591	21.1735	27.6092	27.5918	0.0495	26.1519	27.5369	21.8629	27.587	27.6182	0.0495
	CRC	0.9969	0.9976	0.9893	0.9976	0.9976		0.9967	0.9976	0.991	0.9976	0.9976	
SP (0.1)	PSNR	-5.0013	14.1647	-0.3985	16.0549	16.4447	0.0528	-2.2482	17.8722	1.1564	20.0074	19.8276	0.0508
	CRC	0.2575	0.9487	0.3706	0.9664	0.9682		0.3447	0.9772	0.4384	0.986	0.9854	
SP (0.3)	PSNR	-13.8007	4.9173	-7.5213	6.7054	8.69	0.0803	-10.1943	8.9492	-5.9511	11.3119	12.0809	0.0596
	CRC	0.0729	0.7181	0.1057	0.7902	0.7881		0.1073	0.8531	0.1205	0.908	0.9091	

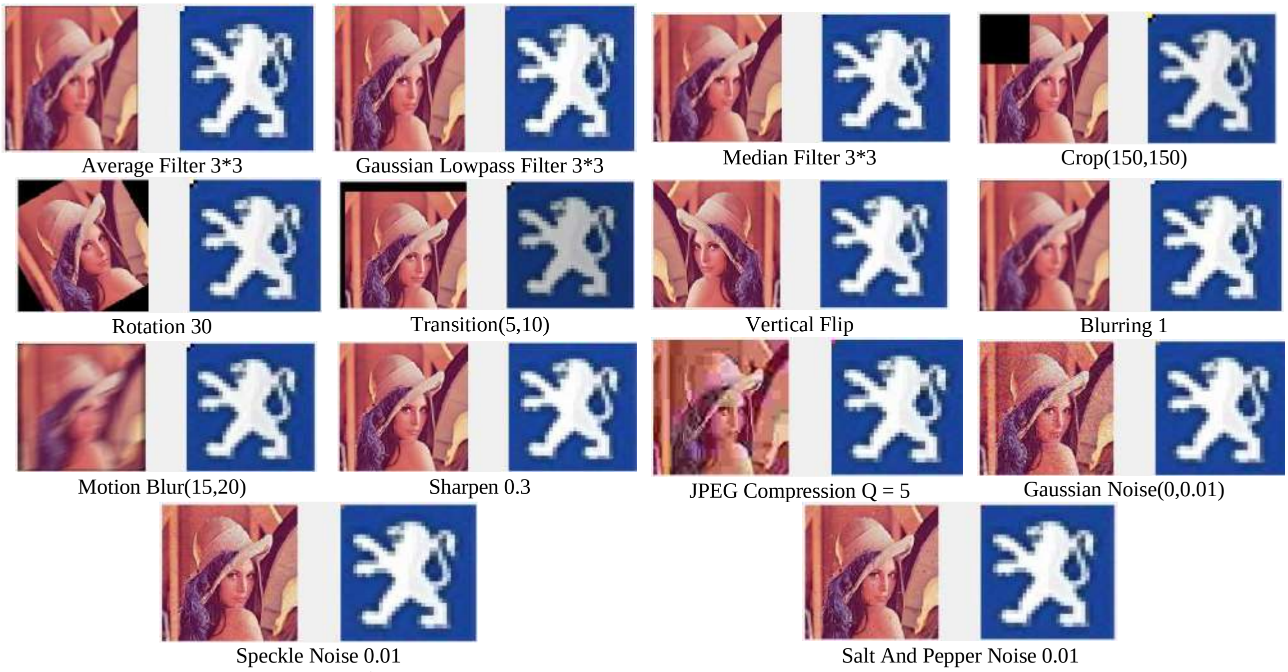


Figure 6. Results of watermarking images (Lena image) and extracted watermark image (Peugeot logo)

We can see the PSNR and the CRC results obtained from the methods tested on different images in Figures 7, 8, 9, and 10, respectively. After various attacks, the result of experiments in the proposed method indicated that we have improved and increased the transparency parameters of PSNR and the CRC robustness parameter over the tested methods. The execution time of the studied algorithms is shown in Table 2.

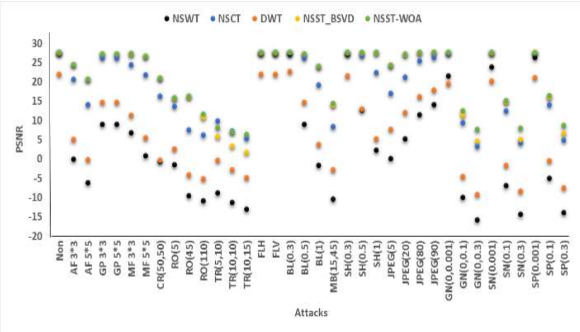


Figure 7. Comparison of the PSNR of the proposed method with the tested methods on the Airplane image

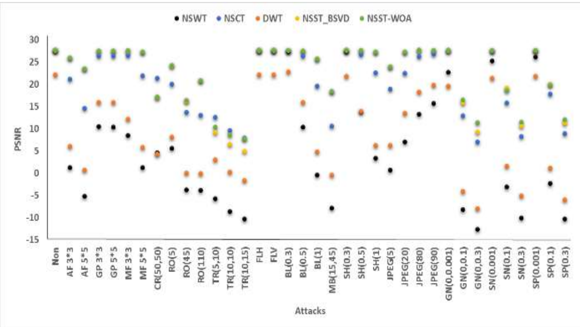


Figure 8. Comparison of the PSNR of the proposed method with the tested methods on the Lena image

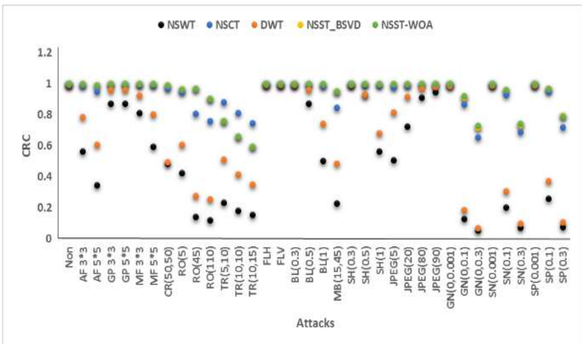


Figure 9. Comparison of the CRC of the proposed method with the tested methods on the Airplane image

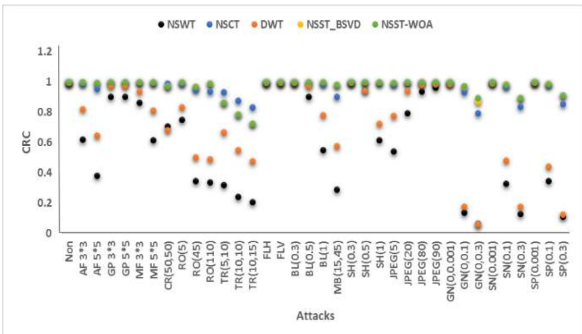


Figure 10. Comparison of the CRC of the proposed method with the tested methods on the Lena image

TABLE 2. Comparison of execution time of the studied algorithms

Algorithm	NSWT	NSCT	DWT SVD	DST BSVD	NSST WOA
Time (second)	0.92	64.23	0.98	2.31	9.90

6. CONCLUSION

In this paper, a new watermarking method using the Shearlet Transform domain coupled with the SVD transform and Whale optimization algorithm is presented with higher transparency and robustness than the comparison methods. Shearlet Transform has more transparency than traditional converts. The SVD transform also increases the robustness of watermarking operations. The whale optimization algorithm is also used to recover the watermark with the least damage in the watermark extraction step. An optimal combination of these methods has resulted in good transparency and resistivity watermarking. The results are extensible in future work and can be tested using other optimization transformations and algorithms.

7. REFERENCES

1. Huang, B.-B. and Tang, S.-X., "A contrast-sensitive visible watermarking scheme", ***IEEE MultiMedia***, Vol. 13, No. 2, (2006), 60-66. DOI: 10.1109/MMUL.2006.23
2. Abdelhakim, A.M., Saleh, H.I. and Nassar, A.M., "A quality guaranteed robust image watermarking optimization with artificial bee colony", ***Expert Systems with Applications***, Vol. 72, No., (2017), 317-326. DOI: 10.1016/j.eswa.2016.10.056
3. Das, C., Panigrahi, S., Sharma, V.K. and Mahapatra, K., "A novel blind robust image watermarking in dct domain using inter-block coefficient correlation", ***AEU-International Journal of Electronics and Communications***, Vol. 68, No. 3, (2014), 244-253. DOI: 10.1016/j.aeue.2013.08.018
4. Wu, X. and Sun, W., "Robust copyright protection scheme for digital images using overlapping dct and svd", ***Applied Soft Computing***, Vol. 13, No. 2, (2013), 1170-1182. DOI: 10.1016/j.asoc.2012.09.028
5. Ali, M., Ahn, C.W. and Siarry, P., "Differential evolution algorithm for the selection of optimal scaling factors in image watermarking", ***Engineering Applications of Artificial Intelligence***, Vol. 31, (2014), 15-26. DOI: 10.1016/j.engappai.2013.07.009
6. Ali, M. and Ahn, C.W., "An optimized watermarking technique based on self-adaptive de in dwt–svd transform domain", ***Signal Processing***, Vol. 94, (2014), 545-556. DOI: 10.1016/j.sigpro.2013.07.024
7. Mishra, A., Agarwal, C., Sharma, A. and Bedi, P., "Optimized gray-scale image watermarking using dwt–svd and firefly algorithm", ***Expert Systems with Applications***, Vol. 41, No. 17, (2014), 7858-7867. DOI: 10.1016/j.eswa.2014.06.011
8. Makbol, N.M., Khoo, B.E. and Rassem, T.H., "Block-based discrete wavelet transform-singular value decomposition image watermarking scheme using human visual system

characteristics", ***IET Image Processing***, Vol. 10, No. 1, (2016), 34-52. DOI: 10.1049/iet-ipr.2014.0965

9. Makbol, N.M. and Khoo, B.E., "A new robust and secure digital image watermarking scheme based on the integer wavelet transform and singular value decomposition", ***Digital Signal Processing***, Vol. 33, (2014), 134-147. DOI: 10.1016/j.dsp.2014.06.012

10. Ansari, I.A., Pant, M. and Ahn, C.W., "Robust and false positive free watermarking in iwt domain using svd and abc", ***Engineering Applications of Artificial Intelligence***, Vol. 49, (2016), 114-125. DOI: 10.1016/j.engappai.2015.12.004

11. Makbol, N.M. and Khoo, B.E., "Robust blind image watermarking scheme based on redundant discrete wavelet transform and singular value decomposition", ***AEU-International Journal of Electronics and Communications***, Vol. 67, No. 2, (2013), 102-112. DOI: 10.1016/j.aeue.2012.06.008

12. Araghi, T.K., Abd Manaf, A. and Araghi, S.K., "A secure blind discrete wavelet transform based watermarking scheme using two-level singular value decomposition", ***Expert Systems with Applications***, Vol. 112, (2018), 208-228. DOI: 10.1016/j.eswa.2018.06.024

13. Araghi, T.K. and Abd Manaf, A., "An enhanced hybrid image watermarking scheme for security of medical and non-medical images based on dwt and 2-d svd", ***Future Generation Computer Systems***, Vol. 101, (2019), 1223-1246. DOI: 10.1016/j.future.2019.07.064

14. Najafi, E. and Loukhaoukha, K., "Hybrid secure and robust image watermarking scheme based on svd and sharp frequency localized contourlet transform", ***Journal of Information Security and Applications***, Vol. 44, (2019), 144-156. DOI: 10.1016/j.jisa.2018.12.002

15. Ali, M., Ahn, C.W., Pant, M. and Siarry, P., "An image watermarking scheme in wavelet domain with optimized compensation of singular value decomposition via artificial bee colony", ***Information Sciences***, Vol. 301, (2015), 44-60. DOI: 10.1016/j.ins.2014.12.042

16. Mardanpour, M. and Chahooki, M.A.Z., "Robust transparent image watermarking with shearlet transform and bidiagonal singular value decomposition", ***AEU-International Journal of Electronics and Communications***, Vol. 70, No. 6, (2016), 790-798. DOI: 10.1016/j.aeue.2016.03.004

17. Fazli, S. and Moeini, M., "A robust image watermarking method based on dwt, dct, and svd using a new technique for correction of main geometric attacks", ***Optik-International Journal for Light and Electron Optics***, Vol. 127, No. 2, (2016), 964-972. DOI: 10.1016/j.ijleo.2015.09.205

18. Vahedi, E., Zoroofi, R.A. and Shiva, M., "Toward a new wavelet-based watermarking approach for color images using bio-inspired optimization principles", ***Digital Signal Processing***, Vol. 22, No. 1, (2012), 153-162. DOI: 10.1016/j.dsp.2011.08.006

19. Maity, S.P., Maity, S., Sil, J. and Delpha, C., "Collusion resilient spread spectrum watermarking in m-band wavelets using ga-fuzzy hybridization", ***Journal of Systems and Software***, Vol. 86, No. 1, (2013), 47-59. DOI: 10.1016/j.jss.2012.06.057

20. Papakostas, G.A., Tsougenis, E. and Koulouriotis, D.E., "Moment-based local image watermarking via genetic optimization", ***Applied Mathematics and Computation***, Vol. 227, (2014), 222-236. DOI: 10.1016/j.amc.2013.11.036

21. Tsai, H.-H., Jhuang, Y.-J. and Lai, Y.-S., "An svd-based image watermarking in wavelet domain using svr and pso", ***Applied Soft Computing***, Vol. 12, No. 8, (2012), 2442-2453. DOI: 10.1016/j.asoc.2012.02.021

22. Tsai, H.-H., Lai, Y.-S. and Lo, S.-C., "A zero-watermark scheme with geometrical invariants using svm and pso against geometrical attacks for image protection", ***Journal of Systems and Software***, Vol. 86, No. 2, (2013), 335-348. DOI: 10.1016/j.jss.2012.08.040

23. Ansari, I.A. and Pant, M., "Multipurpose image watermarking in the domain of dwt based on svd and abc", ***Pattern Recognition Letters***, Vol. 94, (2017), 228-236. DOI: 10.1016/j.patrec.2016.12.010

24. Serdean, C.V., Tomlinson, M., Wade, J. and Ambroze, A.M., "Protecting intellectual rights: Digital watermarking in the wavelet domain", in IEEE Int. Workshop Trends and Recent Achievements in IT. (2002), 16-18.

25. Jian, Z., Shan, L., Jian, J., Wanru, Z. and Shunli, Z., "Extended shearlet-based image watermarking algorithm using selective coefficients in horizontal cone", ***Arabian Journal for Science and Engineering***, Vol. 42, No. 8, (2017), 3597-3607. DOI: 10.1007/s13369-017-2620-1

26. Wang, X.-y., Liu, Y.-n., Xu, H., Wang, A.-l. and Yang, H.-y., "Blind optimum detector for robust image watermarking in nonsubsampled shearlet domain", ***Information Sciences***, Vol. 372, (2016), 634-654. DOI: 10.1016/j.ins.2016.08.076

27. Yahya, A.-N., Jalab, H.A., Wahid, A. and Noor, R.M., "Robust watermarking algorithm for digital images using discrete wavelet and probabilistic neural network", ***Journal of King Saud University-Computer and Information Sciences***, Vol. 27, No. 4, (2015), 393-401. DOI: 10.1016/j.jksuci.2015.02.002

28. Kasmani, S.A. and Naghsh-Nilchi, A.R., "Robust digital image watermarking based on joint dwt-dct", (2009). DOI: 10.4156/jdcta.vol3.issue2.amirgholipour

29. Abdallah, H.A., Hadhoud, M.M., Shaalan, A.A. and El-samie, F.E.A., "Blind wavelet-based image watermarking", ***International Journal of Signal Processing, Image Processing and Pattern Recognition***, Vol. 4, No. 1, (2011). DOI:

30. Zain, J.M. and Clarke, M., "Reversible region of non-interest (roni) watermarking for authentication of dicom images", ***arXiv preprint arXiv:1101.1603***, (2011). DOI: arXiv:1101.1603

31. Banitalebi, A., Nader-Esfahani, S. and Avanaki, A.N., "Robust lsb watermarking optimized for local structural similarity", ***arXiv preprint arXiv:1803.04617***, (2018). DOI: arXiv:1803.04617

32. Wenyin, Z. and Shih, F.Y., "Semi-fragile spatial watermarking based on local binary pattern operators", ***Optics Communications***, Vol. 284, No. 16-17, (2011), 3904-3912. DOI: 10.1016/j.optcom.2011.04.004

33. Ni, Z., Shi, Y.-Q., Ansari, N. and Su, W., "Reversible data hiding", ***IEEE Transactions on circuits and systems for video technology***, Vol. 16, No. 3, (2006), 354-362. DOI: 10.1109/TCSVT.2006.869964

34. Arya, P., Tomar, D.S. and Dubey, D., "A review on different digital watermarking techniques", ***International Journal of Signal Processing, Image Processing and Pattern Recognition***, Vol. 8, No. 10, (2015), 129-136.

35. Kaushik, A.K., "A novel approach for digital watermarking of an image using dft", ***International Journal of Electrical, Electronics & Computer Science Engineering***, Vol. 1, No. 1, (2012), 35-41.

36. Vyas, A., Yu, S. and Paik, J., "Multiscale transforms with application to image processing, Springer, (2018).

37. Labate, D., Lim, W.-Q., Kutyniok, G. and Weiss, G., "Sparse multidimensional representation using shearlets", in Wavelets XI, International Society for Optics and Photonics. Vol. 5914, No. Issue, (2005), 59140U. DOI: 10.1117/12.613494

38. Guo, K. and Labate, D., "Optimally sparse multidimensional representation using shearlets", ***SIAM Journal on Mathematical Analysis***, Vol. 39, No. 1, (2007), 298-318. DOI: 10.1137/060649781

39. Favorskaya, M.N., Jain, L.C. and Savchina, E.I., Perceptually tuned watermarking using non-subsampled shearlet transform, in

Computer vision in control systems-4. 2018, Springer.41-69. DOI: 10.1007/978-3-319-67994-5_3

40. Vaidya, P. and PVSSR, C.M., "A robust semi-blind watermarking for color images based on multiple decompositions", ***Multimedia Tools and Applications***, Vol. 76, No. 24, (2017), 25623-25656. DOI: 10.1007/s11042-017-4532-1.

41. Vali, M.H., Aghagolzadeh, A. and Baleghi, Y., "Optimized watermarking technique using self-adaptive differential evolution based on redundant discrete wavelet transform and singular value decomposition", ***Expert Systems with Applications***, Vol. 114, (2018), 296-312. DOI: 10.1016/j.eswa.2018.07.004

42. Mirjalili, S. and Lewis, A., "The whale optimization algorithm", ***Advances in Engineering Software***, Vol. 95, (2016), 51-67. DOI: 10.1016/j.advengsoft.2016.01.008

43. Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P., "Image quality assessment: From error visibility to structural similarity", ***IEEE Transactions on Image Processing***, Vol. 13, No. 4, (2004), 600-612. DOI: 10.1109/TIP.2003.819861

44. Mousavi, S.M., Naghsh, A. and Abu-Bakar, S., "Watermarking techniques used in medical images: A survey", ***Journal of Digital Imaging***, Vol. 27, No. 6, (2014), 714-729. DOI: 10.1007/s10278-014-9700-5

45. Heylen, K. and Dams, T., "An image watermark tutorial tool using matlab", in Mathematics of Data/Image Pattern Recognition, Compression, and Encryption with Applications XI, International Society for Optics and Photonics. Vol. 7075, (2008), 70750D. DOI: 10.1117/12.793308

46. Jabade, V.S. and Gengaje, D.S.R., "Literature review of wavelet based digital image watermarking techniques", ***International Journal of Computer Applications***, Vol. 31, No. 1, (2011), 28-35.

47. Arora, S.M., "A dwt-svd based robust digital watermarking for digital images", ***Procedia Computer Science***, Vol. 132, (2018), 1441-1448. DOI: 10.1016/j.procs.2018.05.076

Persian Abstract

چکیده

نهان نگاری برای پنهان کردن یا اضافه کردن داده یا فایلی در فایل دیگر, به طوری که فقط افراد آگاه با ابزار لازم بتوانند به آن دست یابند و همچنین یکی از راه‌های حفاظت از داده‌های چندرسانه‌ای در برابر کپی برداری‌های غیرقانونی و توزیع غیرقانونی آن ها است. در روش های قبلی نهان نگاری, تصاویر دیجیتال بیشتر از روش هایی مانند تبدیلات فوریه و تبدیلات موجک و روش‌های ریاضی و آماری دیگری استفاده می‌شد. تبدیل قیچک که نوع خاص و نسبتاً جدیدی می باشد می تواند برای نهان نگاری در تصاویر دیجیتال نیز مورد استفاده قرار گیرد. این روش به دلیل ویژگی‌های خاص خود می‌تواند باعث افزایش اثر و بهره وری در کاربردهایی مانند نهان نگاری تصاویر مورد استفاده قرار گیرد. در این مقاله بر خلاف روش‌های قدیمی برای افزایش شفافیت و مقاومت تصویر نهان نگاری شده, نهان نگاره در زیر باند فرکانس پایین تصویر میزبان جاسازی شده است و از تبدیل SVD به همراه الگوریتم بهینه سازی WOA برای به دست آوردن بهترین فاکتور مقیاس آلفا در مرحله استخراج نهان نگاره پس از اعمال انواع حملات مختلف پردازش تصویری استفاده شده است. همچنین روش مورد مطالعه بر روی تصاویر میزبان و تصاویر نهان نگاره رنگی اجرا شده است. نتایج آزمایشات مختلف نشان دهنده این می‌باشد که روش جدید ارایه شده در این مقاله نسبت به روش‌های نهان نگاری مورد مقایسه از لحاظ مقاومت و شفافیت دارای نتایج بهتری در مقابل اکثر حملات پردازش تصویری می باشد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 4, (April 2021) 854-862

International Journal of Engineering

Journal Homepage: www.ije.ir

Improvement of Multi-agent Routing Guidance with an Intelligent Traffic Light Scheduling and the Ability to Select Intermediate Destinations

M. Habibi[a], A. Broumandnia*[b], A. Harounabadi[c]

[a] *Department of Computer Engineering, Qom Branch, Islamic Azad University, Qom, Iran*
[b] *Department of Computer Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran*
[c] *Department of Computer Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran*

PAPER INFO

Paper history:
Received 29 November 2020
Received in revised form 19 January 2021
Accepted 20 February 2021

Keywords:
Multi-agent Routing
Urban Traffic
Intelligent Traffic Light Scheduling
Intermediate Destinations Selection

ABSTRACT

Controlling traffic congestion and path routing are integral parts of urban development in large cities. Car routing and traffic flow have a direct impact on each other. Therefore, the first step is to determine a criterion for assessing the traffic situation. The type of vehicles should also be considered in routing. Emergency vehicles must arrive at their mission site as soon as possible. Public transportations must also travel according to their plans. Ordinary vehicle drivers can either choose a road as an intermediate destination (place of interest, picking up someone, etc.). In this paper, two new algorithms are proposed to 1) route with intermediate destination selection for ordinary vehicles, and 2) schedule traffic lights to decrease traffic density and routing delay. The first algorithm proposes an agent-based routing model that in addition to finding the least expected travel time (LET) routes, drivers could select a part of the route as intermediate destinations according to their interests, to raise their satisfaction level. The second algorithm considers the density of traffic flow and the presence of emergency vehicles. This algorithm evaluates the status of the traffic flow by the fuzzy logic. The evaluation is conducted by considering traffic flow speed and density. The output of fuzzy logic is used by the Gradational Search Algorithm (GSA). The GSA regards the status of the flow, the priority of the traffic flow, and the distance of the emergency vehicles to the traffic light. The simulation results indicate that the proposed algorithms have better performance.

***doi**: 10.5829/ije.2021.34.04a.11*

NOMENCLATURE

u(x)	A vehicle can enter a street or not	e	Eligibility of members
L(x)	Vehicle length limitation of a street	st_i	The status value of the flow
w(x)	Vehicle weight limitation of a street	**Greek Symbols**	
P_i	Probability value for each road	ρ	The amount of pheromone on each path
$\bar{v}$	The flow speed	α_i	The green interval assigned to the flow
d_i	Distance between the vehicle and the traffic light		

1. INTRODUCTION

Today, traffic is one of the most challenging issues around the world. Heavy traffic is a major problem for many drivers since it causes delay and stress. The economic impact of traffic management is also important, as it increases fuel consumption per capita, loss of time, and so forth. These problems lead to a challenging issue, namely the Vehicle Routing Problem (VRP). Classic VRP consists of assigning routes to a fleet of vehicles to provide services to customers [1]. The classical traveling salesman problem (TSP) can be regarded as VRP. This problem is considered deterministic and static; in other words, all information is previously determined, and during the execution of the routing plan, it remains static [2]. VRPs are usually modeled based on graphs. These graphs are composed of

*Corresponding Author Institutional Email: broumandnia@gmail.com (A. Broumandnia)

Please cite this article as: M. Habibi, A. Broumandnia, A. Harounabadi, Improvement of Multi-agent Routing Guidance with an Intelligent Traffic Light Scheduling and the Ability to Select Intermediate Destinations, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 854-862

vertices (representing intersections) and edges (representing streets). Edges generally contain costs that often specify distances or travel times [1]. The VRP includes the search for a path for *k* vehicles so that each vertex is met exactly once, and the total routing cost (like the total travel time or length of the path) is minimized.

Furthermore, as the world becomes more complex, our ability to understand it should also increase. A complex system consists of several components called a multi-agent system. Multi-agent systems are a new and promising area in the field of distributed artificial intelligence, as well as in traffic light scheduling [3]. If the system contains many agents, then there will be many interactions that making it difficult to predict the behavior of a complex system. The total behavior of these interacting components is nonlinear. Different systems of the real world can be considered complex systems. An example of a complex system is the traffic system. In this system, cars act as entities (agents), and the system is called a “multi-agent system”. When talking about the traffic, the motion of each vehicle and its response to the movement of other vehicles are considered. For example, when a route is blocked due to an accident, the car must first stop and then follow other cars at a low speed. The combination of the behavior of vehicles creates a multi-agent and complex traffic system.

The type of streets network and its characteristics can broadly categorize traffic simulation. The two main categories for simulators are highway environment and urban environment simulators. Furthermore, traffic simulators can be divided into microscopic and macroscopic groups based on the depth of vision. Microscopic simulators model the flow of traffic using complex mathematical models, often derived from fluid dynamics, with all vehicles alike, and they use input and output variables such as velocity, current, and density. These simulators cannot model complex routes, characteristics of precise traffic control, or different driver behaviors [4-6]. Macroscopic simulators are more useful to simulate large traffic systems requiring no detailed modeling such as highway networks. Although this method is fast and can be useful and accurate, it is unsuitable for urban models [5]. In microscopic simulators, interactions are usually controlled by vehicle following and lane changing models [7]. Microscopic simulators can more accurately model the flow of traffic than macroscopic simulators do, owing to the added details of modeling the vehicles [6]. Microscopic simulators are widely used to evaluate control and new traffic management technologies, and to analyze existing traffic operations [4, 8]. Urban traffic networks, with many streets and intersections, are extremely complicated. In most cases, they must handle a large number of vehicles in a small part of the street, which can lead to more density and complexity [9].

Recently, companies such as Google Maps [10] and Waze [11] have developed navigation applications trying to find the fastest navigation. Nevertheless, fastness is not the only driver’s concern. However, these applications try to avoid traffic jams, guide routes with less travel time, and suggest multiple ways to navigate to increase the level of satisfaction, but drivers always do not want to travel fast, rather they want to enjoy the ways. In some cases, there could be roads that drivers like to cross even if they have light traffic or their travel time and distance increase (picking up someone, shopping, a memorable path, ease of driving in the driver’s opinion, beauty of a path, etc). Selecting a part of route could enhance the level of the driver’s satisfaction and increase the joy of travel in using these applications.

Moreover, traffic light scheduling algorithms have a considerable impression on the traffic density. These algorithms consider the expected arrival time to the signalized intersection or the density of competing for traffic flows [12]. Therefore, a well-chosen schedule of traffic lights contributes to the efficient processing of vehicle flows at intersections, thereby raising the average speed and decreasing the time spent by cars in traffic jams [13]. Traffic light scheduling algorithms have to evaluate the conditions of traffic flows. Such an evaluation should be conducted by considering one or more parameters. The results of the evaluations must reflect the current conditions and somewhat should predeict the traffic flow. Scheduling of traffic lights should be carried out dynamically and intelligently. The proposed algorithm aims to increase the speed of flows in streets leading to the junction, and to decrease the waiting time of vehicles at traffic lights. It aims to provide the least possible delay for emergency vehicles (i.e. ambulances, fire engines, and police scouts) and buses. It schedules traffic lights by using the Gravitational Search Algorithm (GSA). It evaluates the conditions of flows using fuzzy logic. This paper, also, introduces a kind of multi-agent vehicular urban traffic guidance that uses the driver’s selection of multiple roads as intemediate destinations. This system includes finding the shortest path with a less travel time algorithm, dynamic rerouting, and considering the driver’s selection for intermediate destinations.

The remainder of this article is organized as follows: Section 2 discusses related works. Section 3 describes the proposed approach. Section 4 provides the simulation and results, and finally, Section 5 presents the conclusion.

2. RELATED WORKS

Multi-agent models have received increasing attention in traffic management, signal control, and route guidance. This method has special benefits such as greater flexibility, faster response, robustness, resource sharing, and better compatibility in aggregating pre-existing and independent systems [14].

Markov and Gray's models predict traffic volume [15]. Genetic algorithms are also used to predict the traffic volume. A comparison has been conducted in literature made to determine the optimal method to predict the short-term traffic volume between the three methods [16, 17]. The source offers model time windows provided by customers [18], which are considered fuzzy random variables. Group intelligence patterns, especially ant colony, have been widely used to solve complex problems. Ant colony algorithm utilizes the behavior of ants and uses the pheromone for routing. There are various conducted researches in this field [19-21].

Recent advances in ICT and accessories, including Global Positioning System (GPS) have focused on deterministic vehicle routing problem (DVRP) In particular, intelligent transport systems (ITS) can collect and process geographic data (such as vehicle position and status, and traffic information) in real-time and design operations [22]. More specific systems, including traffic simulators, are described in litertaures [22, 23].

Most old VRPs rely on computing the least expected travel time (LET) paths for each vehicle to reduce total travel time [24, 25] that is not very useful for urban traffic. Furthermore, dynamic traffic, erratic events, and guidance of drivers toward a similar route lead to unstable global traffic behavior. The last situation could be avoided with dynamic user optimal traffic assignment [26] that periodically computes the assignment of traffic flows to routes.

Vehicle routing algorithms in vehicular traffic guidance systems have been modified several times. However, in recent research [27], arriving on time and total travel time have been considered as two important factors in vehicle routing guidance systems. However, the older systems always consider them separately, since they may have conflicts. This article proposes a semi-decentralized multi-agent system that adaptively provides route guidance at each road intersection by considering the intentions of vehicle agents.

An intelligent trafic routing algorithm was introduced in litertaure [28]. This algorithm considers the real-time traffic flow intending to cross the intersection of interest and schedule traffic light's time phases of each. This algorithm aims to increase traffic fluency by decreasing the waiting time of traveling and to increase the number of vehicles crossing the intersection per second.

Faye et al. introduced an algorithm to exploit Wireless Sensor Networks (WSN) [29]. Sensor nodes are deployed in lanes that sample the number of vehicles passing them. This is used to schedule traffic lights. The sensor nodes are organized in hierarchical order. Each sensor reports to its upper-level sensor. The traffic light scheduling is carried out by the sensor nodes at the third level. Rezgui et al. proposed three different algorithms to schedule traffic lights [30]. These algorithms are developed by using a heuristic method considering the density of flows as well as the waiting time of vehicles. These algorithms are named Smart Traffic Light Scheduling based on traffic density (STLSD), Smart Traffic Light Scheduling based on Density and delay Time (SDLSDT), and the STLSDT that takes into account emergency vehicles (SDLSDE). The STLSD considers the density of flows for scheduling. The SDLSDT regards density and the wait time of vehicles for scheduling. The SDLSDE improves SDLSDT by considering emergency vehicles. Different configurations of a genetic algorithm to find an effective traffic light schedule, which is based on a genetic algorithm by the process of natural selection, are discussed in literature [12].

Sharma and Indu [31] proposed a new traffic light scheduling, which uses GPS information. This algorithm assumes that a significant proportion of people use map services like Google Maps, and they follow the route direction using these services. Razavi et al. proposed an algorithm by combining Internet of Things (IoT), image, and video processing techniques in [32]. It considers the flow density and the number of passing vehicles. This algorithm employes two approaches: 1) uses an image of the traffic conditions to determine the density of flow, and 2) uses live video to count the number of vehicles in streets leading to junctions.

Hu et al. proposed a hybrid cellular swarm optimization method to optimize the scheduling of urban traffic lights [33]. Particle swarm optimization plays a pivotal role in this framework which exploits two models: Phase scheduling model (PSM) and Phase timing scheduling optimization model (POM). The PSM conserves the scheduling settings and applies transition rules to achieve comprehensive scheduling. The POM is devoted to the optimization of the phase timing schedule.

The main contributions of this paper are summarized as follow:

- Determine a criterion for assessing the traffic situation
- Consider the type of vehicles in routing
- Introduce an algorithm to route with intermediate destination selection for ordinary vehicles
- Introduce of an algorithm to schedule traffic lights to decrease traffic density and routing delay
- Evaluates the status of the traffic flow by fuzzy logic
- Regards status of the flow, the priority of the traffic flow, and the distance of the emergency vehicles to the traffic light with the GSA method

3. PROPOSED APPROACH

The proposed vehicle routing solves the complicated routing problems in three levels to make it a simpler problem. The three levels are local, regional, and urban levels. This kind of partitioning allows solving the routing problem part by part. A grid is laid out in Tehran city. The entire city surface is related to the urban level. The cells in Figure 1 represent the regional level, and the cells in Figure 2 represent the local level partitioning. An agent is placed for each cell at local and regional levels. An agent is also placed at the urban level that is the central agent. The map of city is available for all agents.

A way to evaluate the traffic conditions of each road in the city should be devised. In this paper, the fourth agent is proposed to count the number of vehicles and the speed of traffic flows on each road. The proposed approach needs to place another agent on each road. This agent reports only the number of vehicles and the speed of each traffic flow. These reports are sent to the local agents. The agents evaluate the conditions of each road and report the overall conditions of their cell.

It is obvious that any type of vehicle cannot enter any road. Some roads have traffic limitations. For example, there may be a bridge in the road that cannot bear vehicles with a weight of more than 10 tonnes. These limitations should be considered in the routing. In this paper, we consider only the size and weight of vehicles. Equation (1) evaluates whether a vehicle can enter a street or not. Equation (2) checks the length limitation of the street $road_l$ for the requested vehicle $vehicle_l$. Equation (3) checks the weight limitation of the street $road_w$ for the vehicle $vehicle_w$.

$$u(x) = \begin{cases} 1 & L(x) = 1 \wedge w(x) = 1 \\ 0 & L(x) = 0 \vee w(x) = 0 \end{cases} \tag{1}$$

$$L(x) = \begin{cases} 1 & road_l \geq vehicle_l \\ 0 & road_l < vehicle_l \end{cases} \tag{2}$$

$$w(x) = \begin{cases} 1 & road_w \geq vehicle_w \\ 0 & road_w < vehicle_w \end{cases} \tag{3}$$

3. 1. Routing at the Local Level In the proposed algorithm, the driver sets his/her destination by searching and pinning on the map. The driver is allowed to select a road that he or she wants to pass along the road. This information and the vehicle weight, length, and width are sent to the local agent. Table 1 shows the request format. This request may be sent by the vehicle or the regional agent of the local agent. Each local agent is supervised by a regional agent. All regional agents are supervised by the urban agent.

The map of the city is converted to a directed graph. Each branch in the network of roads is represented by a node. Each node has a label. The information embedded in the label includes the identification (ID) and the coordination of that branch. Each road is represented by an edge. These edges are labeled by the ID, length, traffic flow density, traffic flow speed, and traffic limitations of the road. The local agents use this graph in their routing process.

The local agent locates the source and destination coordination on the map and identifies the ID of the road on which the source and the destination points are placed. The agent should determine whether the routing is within the borders of its cell or not. If the routing is beyond its borders, it forwards the routing request to its regional agent. The regional agent will do the same process. In other words, if the routing is beyond the regional cell, then the urban agent should start the routing process. Figure 3 presents the algorithm.

Let suppose that the local agent has received a routing request whether from the vehicle or the regional agent. This agent extracts the vehicle specifications and its driver's demands. It starts a routing process based on the

TABLE 1. The request format for a vehicle

Requested formats		
Id	Source Coords	Destination Coords
Length	Width	Height
Desired Strret Id	Arrival time	Tolerable delay

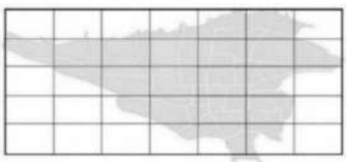


Figure 1. Regional partitioning on Tehran city

Figure 2. Local portioning

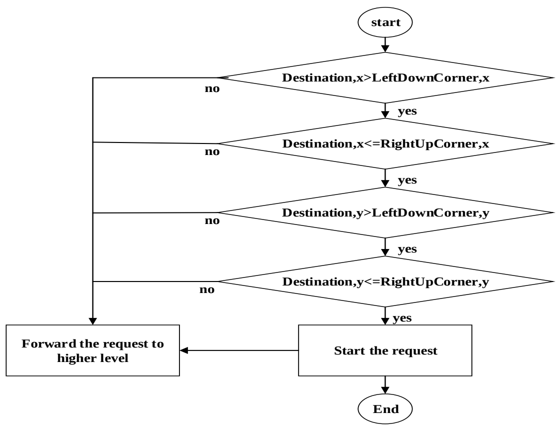


Figure 3. Algorithm of routing at local or higher level

ant colony optimization. It first creates a sufficient number of ants with a label including the vehicle specifications and demands. These ants move in the direction of the flow traffic in which the vehicle is included. When each ant meets a branch, it will calculate a probability value for each road. These values are called the probability selection value. It will, then, define a selection interval for each road. These intervals are used to select the roads randomly. The selection probability value is calculated based on Equation (4).

$$P_i = \begin{cases} \frac{flow\ speed}{\max Speed + density} & u(x) = 1 \\ 0 & u(x) = 0 \end{cases} \tag{4}$$

The selection interval values are defined based on the algorithm shown in Figure 4. The ant will generate a random number in the interval [0,1]. Each interval is assigned to a road. The road will be selected only if this number falls within its selection interval. The union of intervals is the interval [0,1]. Longer intervals are likely selected. According to the algorithm in Figure 4, those roads having higher selection probability value will have longer selection intervals.

The ants will pass the selected road. The aim of the algorithm in Figure 4 is to discover all the possible paths between the source and the destination. The ants will return to the source only if they meet the destination. The ants will record the ID of roads that they pass. No ant will pass a road twice. If an ant meets a branch or road twice, it will die. The ants have a lifetime so that if they do not meet the destination in their lifetime, they will die. In other words, those ants are lost. If the driver demands to pass a certain road and an ant meets a destination but does not pass the demanded road, the ant will die.

The ants that have succeeded in meeting the destination will return to the source by the path that they have passed along it. They will lay some pheromone out according to Equation (5). The Equation calculates the pheromone-based on the reported traffic conditions of the

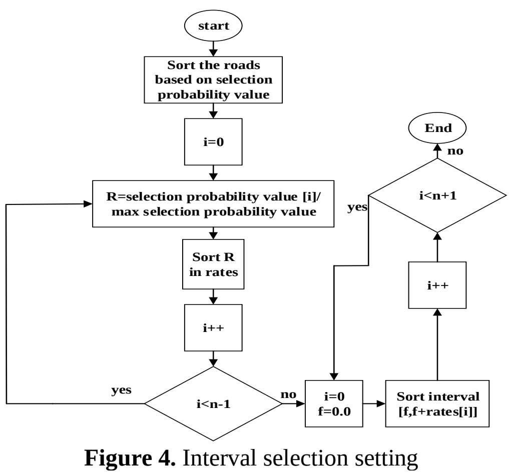


Figure 4. Interval selection setting

road and the driver's demands, where the $road_i$ is the length of ith road along the discovered route.

$$\rho = \begin{cases} \left(\frac{\sum_{i \in path} road_i}{flowSpeed} - \frac{\sum_{i \in path} road_i}{maxSpeed}\right), Delay_{real} < Delay_{deamanded} \\ \frac{flowSpeed}{maxSpeed + flowDensity + \sum_{i \in path} road_i}, \ Delay_{real} \geq Delay_{deamanded} \end{cases} \tag{5}$$

3. 2. Routing at Regional and City Level

As mentioned earlier, when a local agent receives a routing request, it will forward the request to its regional agent. The regional agent evaluates the request to determine whether it is within its borders or not. If the request is beyond its borders, it will forward the request to the urban agent. Urban and regional agents will run the same algorithm with a slight difference.

Each upper-level agent will evaluate the conditions of its scope by the report of lower-level agents. Local agents evaluate the conditions of its scope by the reports of traffic control agents. Regional agents evaluate its region by the reports of local agents. Urban level agents evaluate the conditions of the city through the reports provided by regional agents. The agents at these levels perform the same algorithm.

The vehicle routing in these two levels is as simple as drawing a straight line between two points. The two points are the source and the destination of a vehicle. This line will determine local and regional cells that should be activated for routing. The city agent determines regional cells that should go through the routing process. If the traffic condition in a regional cell is more than the threshold, the urban routing agent will replace the cell with its neighbor regional cells. This replacement will be done if the traffic condition of the neighbor regional cell is more than the threshold, otherwise, there will be no replacement.

The urban agent will distribute the driver's demands between the regional agents that should run the routing algorithm. The regional agent will distribute the driver's demands between the local agents. Suppose that a driver has demanded a route to a location at 20 O'clock in such a way that he or she wants to be at the destination in 22 O'clock with a tolerable delay of 30 minutes. Suppose that there are four regional cells between the source and the destination. The urban agent will request the first regional agent to route the vehicle to the out of its borders until 20:30 with a delay of 12 minutes. The other regional agents have to route the vehicle out of their borders at 21:00, 21:30, and 22:00, respectively.

3. 3. Traffic Control Using Intelligent Traffic Lights

Traffic lights are mostly located at intersections. In any intersection, there are usually four incoming traffic flows and eight outgoing traffic flows. This algorithm [34] assigns an integer number between 1 and 8 to each outgoing flow and evaluates incoming traffic flows by considering the speed and density of the flow as well as the distance of the emergency vehicle to the traffic light. Four types of emergency vehicles are considered: ambulance, fire engine, police scouts, and busses. A priority is an integer between 1 and 5 to each flow. The fuzzy engine aims to evaluate the status of each incoming flow to a junction. The evaluations are conducted by two parameters, namely speed, and density of the flow. The outcome of this test is the ratio of vehicles that can pass the traffic light. If the outcome is zero, no vehicle can pass and if the outcome is one, all vehicles can pass the junction. Figure 5 illustrates the fuzzy system.

3. 3. Scheduling Traffic Lights Our proposed scheduling [34] is carried out based on the passing ratio, the priority of the flow, and the distance of the emergency vehicle to the junction. Total green intervals in each junction fall into [min, max]. In other words, the i total green interval could be 3 minutes, and the i+1 is 2.5 minutes. This allows the emergency vehicles, which are close to the traffic light, to pass immediately, and another flow receives the green light. Scheduling is performed by the gravitational search algorithm. The initial population is produced by the algorithm in Figure 6.

The members of the initial population are created randomly. The population for each traffic light may differ

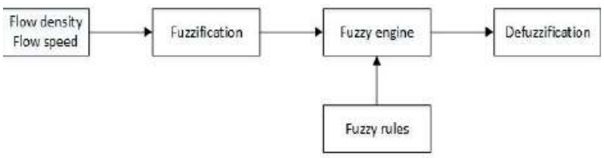


Figure 5. Block diagram of the fuzzy system for vehicle routing.

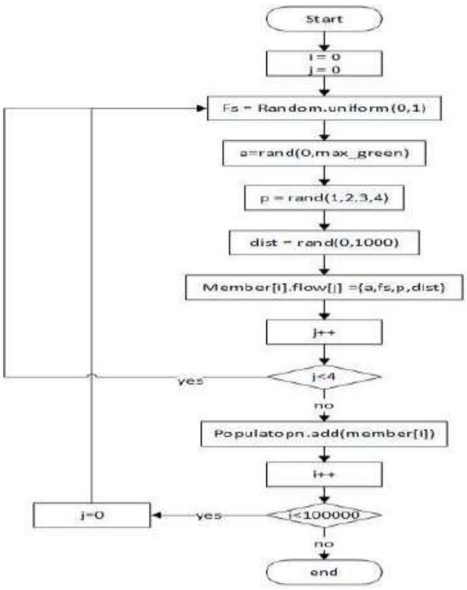


Figure 6. Initial population production

and is optimized gradually. Each member has four flows and an eligibility value, which is determined using Equation (6), where st_i is the status value of the flow, p_i is the priority of the flow, d_i is the distance of the emergency vehicle to the traffic light, and α_i is the green interval assigned to the flow.

$$A = \pi r^2 = \frac{\alpha_1 \times p_1 \times d_1}{st_1} + \frac{\alpha_2 \times p_2 \times d_2}{st_2} + \frac{\alpha_3 \times p_3 \times d_3}{st_3} + \frac{\alpha_4 \times p_4 \times d_4}{st_4} \quad (6)$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$$

4. SIMULATION AND RESULTS

The simulations are run on a platform programed in C#. In this platform, the unit of time is second (s), the length is in meter (m) and the speed is meter per second (m/s). A total of 16 regional cells are defined and each regional cell is further divided into 16 local cells. To evaluate the proposed algorithm, three different scenarios are run. The difference between these scenarios is the number of vehicles. There are 144 intersections in the scenarios. In the meantime, the proposed algorithm by Razavi et al. [33] is considered as the base algorithm and have considerable results. Table 2 presents the conditions of the simulation.

4. 1. The Average Speed of Vehicles Average speed of vehicles is considered to evaluate the fluency of flows on intersections. This parameter is calculated by averaging the average speeds of all vehicles that traveled to their destination. As Figure 7 shows, the results of average speed for 2500, 5000 and 10000 vehicles determine that the proposed algorithm has better performance.

4. 2. Average Delay This parameter shows the average delay of vehicles in intersections. The delay interval of vehicles in the control area of traffic lights is recorded and is averaged at the end of the simulation. The simulation results in Figure 8 indicate that the proposed algorithm had a shorter delay.

TABLE 2. Simulation Conditions

Items	Value		
Junctions	144		
Junctions type	Grid		
Number of ordinary cars	2500	5000	10,000
Number of emergency vehicles (ambulance, fire, and police)	100	200	300
Simulation time	1 hour		

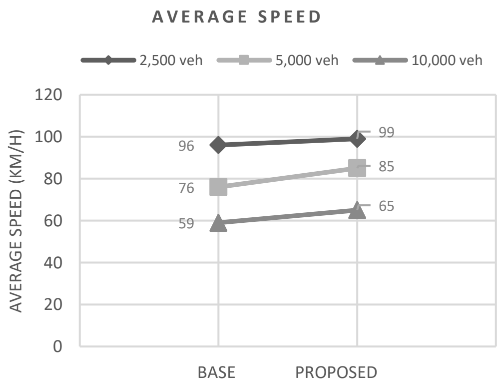


Figure 7. Average speed by the 2500, 5000 and 10000 vehicles.

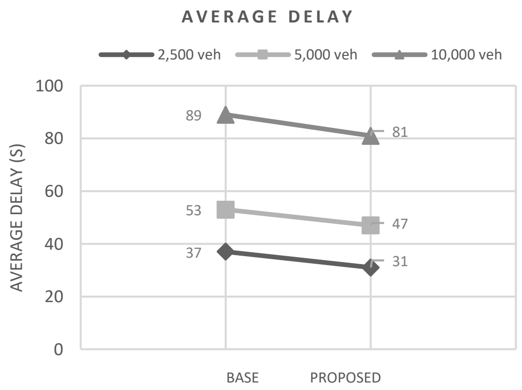


Figure 8. Average delay by the 2500, 5000 and 10000 vehicles

4. 3. Arrival Time

The proposed algorithm has considered the density of each road and the speed of flow in each road. This strategy resulted in routing with more reliability and had a better prediction.

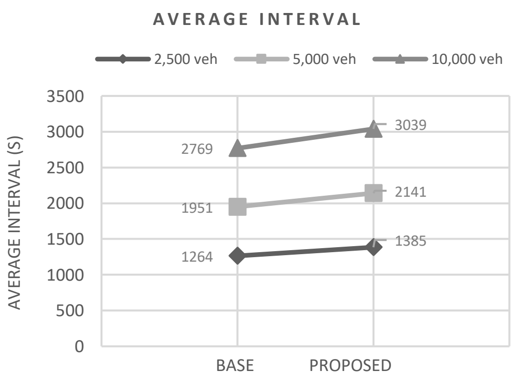


Figure 9. Average arrival time for the 2500, 5000 and 10000 vehicles.

According to Figure 9, the vehicles have averagely arrived at their destination approximately 120 seconds earlier in the 2500 vehicle scenario, 190 seconds earlier in the 5000 vehicle scenario, and 270 seconds earlier in the 10000 vehicle scenarios. As can see, the results are better than the results of the base algorithm.

4. 4. Total Travel Time

The focus of this research was to route the vehicles in such a way that they arrive at their destination on time. To accomplish this, the total travel time should be reduced. The proposed algorithm, therefore, considered the density and speed of flows. Thus, the discovered paths are faster and more reliable. Furthermore, the paths discovered by the ant colony optimization tend to be the shortest ones, since this optimization method tends to find the shortest path. However, the shortest paths are not always the best paths. This tendency is controlled by Equation (6). The amount of pheromone on each path is a function of the driver's demands. According to Figure 10, the total travel time of 2500, 5000 and 10000 vehicles was respectively 378s, 466s, and 694s earlier under the routing of the proposed algorithm.

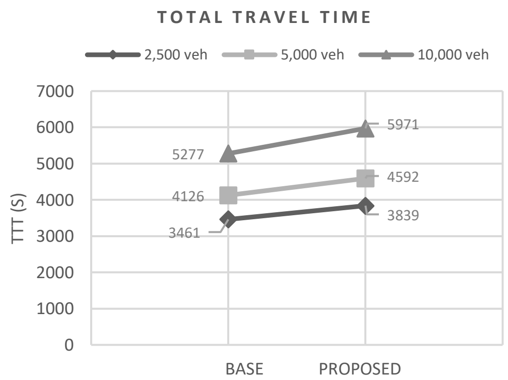


Figure 10. TTT of the 2500, 5000, and 10000 vehicles

5. CONCLUSION

Traffic congestion, transportation, and route guidance are integral development parts of large cities. Route guidance aims to find the fastest path from the origin to the destination. This is while drivers may be interested in passing along a particular road as an intermediate destination (picking up someone else, shopping, a memorable path, ease of driving in the driver's opinion, beauty of a path, etc.).

This paper provides an intelligent solution to control and manage the crossroad traffic using the fuzzy logic and gravitational search algorithm. Scheduling of green and red traffic lights is determined based on the density and speed of the flows passing through the main streets leading to the junctions. The proposed algorithm has provided less delay, average arrival time, and total travel time for vehicles. It has also considered emergency vehicles and busses. It performed more efficiently than the base algorithm. In future studies, the proposed algorithm will be tested for different penetration rates, and its behavior will be investigated in the case that all vehicles are not equipped with vehicular transceivers.

5. REFERENCES

1. Ulmer, M.W., Approximate dynamic programming for dynamic vehicle routing, Springer, 2017.
2. Helsgaun, K., "An extension of the lin-kernighan-helsgaun tsp solver for constrained traveling salesman and vehicle routing problems", ***Technical Report,*** *Roskilde: Roskilde University*. (2017).
3. Badjonski, M., Ivanović, M. and Budimac, Z., Agent oriented programming language lass, Object-oriented technology and computing systems re-engineering, 1999.
4. Ferber, J. and Weiss, G., Multi-agent systems: An introduction to distributed artificial intelligence, Addison-Wesley Reading, Vol. 1, 1999.
5. Lemessi, M., "An slx-based microsimulation model for a two-lane road section", in Proceeding of the 2001 Winter Simulation Conference (Cat. No. 01CH37304), IEEE. Vol. 2, 1064-1071, (2001), doi: 10.1109/WSC.2001.977415.
6. Owen, L.E., Zhang, Y., Rao, L. and McHale, G., "Traffic flow simulation using corsim", in 2000 Winter Simulation Conference Proceedings, IEEE. Vol. 2, Issue, 1143-1147, (2000), doi: 10.1109/WSC.2000.899077.
7. Schulze, T. and Fliess, T., "Urban traffic simulation with psycho-physical vehicle-following models", Proceedings of the 29th conference on Winter simulation. 1222-1229, (1997), doi: 10.1145/268437.268764.
8. Lansdowne, A., Traffic simulation using agent-based modelling, University of the West England Bristol, (2006).
9. Nagel, K. and Schreckenberg, M., "A cellular automaton model for freeway traffic", ***Journal de Physique I***, Vol. 2, No. 12, (1992), 2221-2229, doi:10.1051/jp1:1992277.
10. *Google apps*. 2021, URL: https://www.google.com/maps.
11. Waze, *Free community-based gps, maps & traffic navigation app*. 2021, URL: https://www.waze.com/.
12. Younes, M.B. and Boukerche, A., "An efficient dynamic traffic light scheduling algorithm considering emergency vehicles for intelligent transportation systems", ***Wireless Networks***, Vol. 24, No. 7, (2018), 2451-2463, doi: https://doi.org/10.1007/s11276-017-1482-5.
13. Davydov, I., Tolstykh, D., Kononova, P. and Legkih, I., "Genetic based approach for novosibirsk traffic light scheduling", in 2019 15th International Asian School-Seminar Optimization Problems of Complex Systems (OPCS), IEEE. Vol., No. 31-36, (2019), doi: 10.1109/OPCS.2019.8880158.
14. Wahle, J., Bazzan, A.L.C., Klügl, F. and Schreckenberg, M., "The impact of real-time information in a two-route scenario using agent-based simulation", ***Transportation Research Part C: Emerging Technologies***, Vol. 10, No. 5-6, (2002), 399-417, doi: https://doi.org/10.1016/S0968-090X(02)00031-1.
15. Gao, B., Zhang, Q.-Y., Liang, Y.-S., Liu, N.-N., Huang, C.-B. and Zhang, N.-T., "Predicting self-similar networking traffic based on emd and arma", ***Journal of China Institute of Communications***, Vol. 32, No. 4, (2011), 47-56.
16. Bie, Y., Yang, M. and Pei, Y., "Development of short-term traffic volume prediction models for adaptive traffic control", in International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2015), (2015).
17. Ma, X., Yu, H., Wang, Y. and Wang, Y., "Large-scale transportation network congestion evolution prediction using deep learning theory", ***PloS One***, Vol. 10, No. 3, (2015), e0119044, doi: 10.1371/journal.pone.0119044.
18. Ma, Y., Feng, C., Zhang, J. and Yan, F., "Judging customer satisfaction by considering fuzzy random time windows in vehicle routing problems", in International Conference on Management Science and Engineering Management, Springer. 204-211, (2017), doi: https://doi.org/10.1007/978-3-319-59280-0_16.
19. Ando, Y., Masutani, O., Sasaki, H., Iwasaki, H., Fukazawa, Y. and Honiden, S., "Pheromone model: Application to traffic congestion prediction", in International Workshop on Engineering Self-Organising Applications, Springer, 182-196, (2005), doi: https://doi.org/10.1007/11734697_14.
20. Jiang, S., Zhang, J. and Ong, Y.-S., "A multiagent evolutionary framework based on trust for multiobjective optimization", ***in AAMAS***. , 299-306, (2012).
21. Kurihara, S., Tamaki, H., Numao, M., Yano, J., Kagawa, K. and Morita, T., "Traffic congestion forecasting based on pheromone communication model for intelligent transport systems", in 2009 IEEE Congress on Evolutionary Computation, IEEE. 2879-2884, (2009), doi: 10.1109/CEC.2009.4983304.
22. AbdAllah, A.M.F., Essam, D.L. and Sarker, R.A., "On solving periodic re-optimization dynamic vehicle routing problems", ***Applied Soft Computing***, Vol. 55, (2017), 1-12, doi: https://doi.org/10.1016/j.asoc.2017.01.047.
23. Grzybowska, H. and Barceló, J., "Decision support system for real-time urban freight management", ***Procedia-Social and Behavioral Sciences***, Vol. 39, (2012), 712-725, doi: https://doi.org/10.1016/j.sbspro.2012.03.142.
24. Li, J., Wu, Q. and Zhu, D., "Route guidance mechanism with centralized information control in large-scale crowd's activities", in 2009 International Joint Conference on Artificial Intelligence, IEEE. 7-11, (2009), doi: 10.1109/JCAI.2009.41.
25. Yamashita, T., Izumi, K., Kurumatani, K. and Nakashima, H., "Smooth traffic flow with a cooperative car navigation system", in Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems. 478-485, (2005), doi:10.1145/1082473.1082546.
26. Chiu, Y.-C., Bottom, J., Mahut, M., Paz, A., Balakrishna, R., Waller, S. and Hicks, J., Dynamic traffic assignment: A primer (transportation research circular e-c153), 2011.
27. Cao, Z., Guo, H. and Zhang, J., "A multiagent-based approach for vehicle routing by considering both arriving on time and total travel time", ***ACM Transactions on Intelligent Systems and Technology (TIST)***, Vol. 9, No. 3, (2017), 1-21, doi: https://doi.org/10.1145/3078847.
28. Younes, M.B. and Boukerche, A., "An intelligent traffic light scheduling algorithm through vanets", in 39th Annual IEEE Conference on Local Computer Networks Workshops, IEEE. 637-642, (2014), doi: 10.1109/LCNW.2014.6927714.
29. Faye, S., Chaudet, C. and Demeure, I., "A distributed algorithm for adaptive traffic lights control", in 2012 15th International IEEE Conference on Intelligent Transportation Systems, IEEE, 1572-1577, (2012), doi: 10.1109/ITSC.2012.6338671.
30. Rezgui, J., Barri, M. and Gayta, R., "Smart traffic light scheduling algorithms", in 2019 International Conference on Smart

Applications, Communications and Networking (SmartNets), IEEE. 1-7, (2019), doi: 10.1109/SmartNets48225.2019.9069760.

31. Sharma, K.K. and Indu, S., "Gps based adaptive traffic light timings and lane scheduling", in 2019 IEEE Intelligent Transportation Systems Conference, IEEE. 4267-4274, (2019), doi:10.1109/ITSC.2019.8917153.

32. Razavi, M., Hamidkhani, M. and Sadeghi, R., "Smart traffic light scheduling in smart city using image and video processing", in 2019 3rd International Conference on Internet of Things and Applications (IoT), IEEE, 1-4, (2019), doi: 10.1109/IICITA.2019.8808836.

33. Hu, W., Wang, H., Yan, L. and Du, B., "A hybrid cellular swarm optimization method for traffic-light scheduling", ***Chinese Journal of Electronics***, Vol. 27, No. 3, (2018), 611-616, doi: 10.1049/cje.2018.02.002.

34. Habibi, M., Broumandnia, A. and Harounabadi, A., "An intelligent traffic light scheduling algorithm by using fuzzy logic and gravitational search algorithm and considering emergency vehicles", ***International Journal of Nonlinear Analysis and Applications***, Vol. 11, Special Issue, (2020), 475-482, doi: 10.22075/IJNAA.2020.4706.

Persian Abstract

چکیده

تراکم ترافیک، حمل و نقل و راهنمایی مسیر از اجزای اساسی در توسعه شهری در شهرهای بزرگ است. نحوه مسیریابی خودروها در شهرها و ترافیکی که در هر جاده جریان دارد تأثیر مستقیم بر یکدیگر دارند. بنابراین، اولین قدم تعیین معیاری برای ارزیابی وضعیت ترافیک در هر جاده است. نوع وسایل نقلیه نیز باید در مسیریابی در نظر گرفته شود. وسایل نقلیه اورژانس باید در اسرع وقت به محل ماموریت خود برسند. حمل و نقل عمومی نیز باید طبق برنامه ریزی های خود پیش برود. رانندگان وسایل نقلیه معمولی می توانند خیابانی را به عنوان مقصد میانی انتخاب کنند (از روی علاقه، سوارکردن شخصی و غیره). در این مقاله، دو الگوریتم جدید برای 1) مسیریابی با انتخاب مقصد میانی برای عامل های وسایل نقلیه عادی، و 2) برنامه ریزی چراغ های راهنمایی برای کاهش تراکم ترافیک و تاخیر در مسیریابی پیشنهاد شده است. اولین الگوریتم یک مدل راهنمای مسیر مبتنی بر عامل را پیشنهاد داده است که علاوه بر یافتن مسیرهایی با کمترین زمان سفر (LET)، رانندگان می توانند با توجه به علایق خود، بخشی از مسیر را به عنوان مقصد میانی انتخاب کنند تا سطح رضایت آنان افزایش یابد. الگوریتم دوم تراکم جریان ترافیک و حضور عامل های خودروهای اورژانس را در نظر می‌گیرد. این الگوریتم وضعیت جریان ترافیک را با منطق فازی ارزیابی می کند. ارزیابی با در نظر گرفتن سرعت و تراکم جریان ترافیک انجام می شود. خروجی منطق فازی توسط الگوریتم جستجوی تدریجی (GSA) استفاده می شود. GSA وضعیت جریان، اولویت جریان ترافیک و فاصله وسیله نقلیه اضطراری تا چراغ راهنما را در نظر می‌گیرد. نتایج شبیه سازی نشان می‌دهد که الگوریتم های پیشنهادی عملکرد بهتری دارند.

IJE TRANSACTIONS A:Basics Vol. 34, No. 04, (April 2021) 863-872

International Journal of Engineering

Journal Homepage: www.ije.ir

Private Trajectory Intersection Testing: Is Garbled Circuit Better than Custom Protocols?

M. Dehghan*[a], B. Sadeghiyan[a], E. Khosravian[b]

[a] *Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran*
[b] *Department of Mechanical Engineering, Payame Noor University, Tehran, Iran*

PAPER INFO

Paper history:
Received 14 July 2020
Received in revised form 07 January 2021
Accepted 14 January 2021

Keywords:
Distance Computation
Euclidean Distance
Garbled Circuit
Gröbner Basis
Private Trajectory Intersection

ABSTRACT

In this paper, two protocols are presented for private intersection detection of two moving objects' trajectories. To design the first protocol, we simplify the problem of finding the intersection points to the problem of finding the common roots of the polynomials, which represent the moving objects' trajectories. Thereafter, Gröbner Basis is used to design a novel secure protocol to find the common roots of the polynomials. Another protocol is also designed based on the distance computation of two trajectories' curves. The complexity of the Gröbner-based protocol for finding the common roots of polynomials is numerical. Its complexity is much lower than the complexity of the garbled circuit-based protocol for Euclidean Distance Computation of l points and the complexity of the protocols for private proximity testing.

doi: *10.5829/ije.2021.34.04a.12*

1. INTRODUCTION

Private Trajectory Intersection Testing (PTIT) can be considered as a problem in which some moving objects wish to detect the intersection of their trajectories. However, they intend to keep their trajectories secret, namely, to detect whether the distance between moving objects' trajectories is larger than d (d is a predefined threshold), while the parties do not reveal their trajectories of movement.

The problem of PTIT can have some applications in many scenarios, such as urban traffic management, RoboCup competitions, aircraft ad hoc networks, mobile networks and etc. Suppose that the drivers wish to manage traffic in a way to prevent congestion and distribute traffic in streets and highways without revealing their trajectories. Therefore, they want to use a PTIT Protocol. In addition, in RoboCup competitions, the participants tend to select the best trajectories for their robots and do not collide with other robots. These participants apply a PTIT protocol without revealing their trajectories.

The "privacy" is generally in conflict with collecting, storing, using, processing and sharing of personally identifiable data. The primary objective of privacy measures is to ensure proper protection of private data in the course of processing or dissemination of sensitive information [1]. The main feature of PTIT protocol is privacy preserving of moving objects' trajectories.

Secure multiparty computation is an approach to support privacy. In this approach, a set of parties with secret inputs wish to compute some joint functions of their inputs, while they wish to preserve the privacy property. The first general solution for the problem of secure two-party computation in the presence of semi-honest adversaries was presented by Yao [2]. Later, solutions were provided for the multi-party and malicious adversarial cases by Goldreich et al. [3].

In 2016, Hemenway et al. computed the probability of intersection between satellites' trajectories [4]. In their work, the approach is based on garbled circuits, where the garbled circuit for detection of intersection is so large and has not good performance. Atallah and Du

*Corresponding Author Institutional Email: *motahareh479@aut.ac.ir* (M. Dehghan)

Please cite this article as: M. Dehghan, B. Sadeghiyan, E. Khosravian, Private Trajectory Intersection Testing: Is Garbled Circuit Better than Custom Protocols? International Journal of Engineering, Transactions A:Basics Vol. 34, No. 04, (2021) 863-872

[5] presented secure protocols for computational geometry problems. One of the protocols presented in the work of Atallah and Du is for obtaining the intersection of polygons securely. We may use that protocol for obtaining the intersection of moving objects in a certain time, but using it for obtaining the intersection throughout the trajectories is not efficient.

In addition, Frikken and Atallah [6] presented a protocol for two moving objects. In their work, the paths of moving objects are divided into some line segments. Then, in each moment, the intersection of two line segments is obtained. The protocol presented in the work of Frikken and Atallah is just for two-dimensional space.

Moreover, the problem of Private Proximity Testing (PPT) is a similar issue, which is only applicable in smart devices. It has attracted the attention of some researchers since 2006 [7–23]. In PPT, the proximity of two Location Based Services (LBS) are securely examined. Such protocols allow two parties with smart devices, which have Location Based Services to discover if they are close to each other in physical location, without revealing their individual locations to each other. This problem is similar to the problem of PTIT, but in PPT, the parties should have smart devices with LBS.

The previous researches in literature [7–23] are in three categories. In the first category, the parties send securely their location tags to trusted third party, and the trusted third party evaluates if they are neighbor to each other. The second category is based on symmetric key cryptography, and the last is based on public-key infrastructure.

The protocols in the last category are based on modular exponentiation. If we use of RSA for modular exponentiation, we need 'and' operations n^6 times, 'or' operations $n^3 + n$ times. If we want to apply these protocols for PTIT, we should use them for all the points of trajectories, i.e. we need the modular exponentiation $\alpha.n$ times, where α is a constant number.

In this paper, we present two protocols for PTIT of moving objects. In this regard, two protocols are presented based on Gröbner Basis and Distance Computation of two trajectories, which there is no limitation for space dimension. The Gröbner- based protocol was proposed in [24] by Dehghan and Sadeghiyan. We only present an overview of that protocol. Moreover, we analyze its complexity and compare it with the complexity of the garbled circuit-based protocol for Euclidean Distance Computation of l points.

The approach proposed by Yao [2] is a ground-breaking result, which essentially began the field of secure multiparty computation and is served as the basis for countless papers. In addition to its fundamental theoretic contribution, Yao's protocol is remarkably efficient in that it has only a constant number of rounds and uses one oblivious transfer per input bit only (with no additional oblivious transfers in the rest of the computation). However, in the problems with large inputs and circuit size, it might be not efficient.

One way to analyze the complexity of secure multi-party computation protocols, beside the comparison to previous related proposed protocols, is to compare the complexity of the proposed protocols with those of the Yao's garbled circuit-based protocol [2], which is the first general solution.

In this regard, basic operations are implemented using the Yao's garbled circuit [25, 26] (regardless of the proposed security approach) and its computation and communication complexities are examined and compared with the secure proposed protocol.

The basic operation for finding the Intersection of two moving objects' trajectories regardless security requirement can be considered in such a way that each moving object has l discrete points.

Since in Yao's garbled circuit, the circuit gates and their truth tables are formed based on the number of input bits and basic operations, it is necessary to compare the complexity of secure proposed protocols and garbled circuit-based protocols upon the number of input bits and basic operations, as well.

In this paper, we present two protocols for PTIT in Section 2. Then, in Section 3 we prove the security of our proposed protocol based on ideal- real simulation paradigm. We also analyze and compare the computation and communication complexities of our Gröbner-based protocol and garbled cuircuit-based protocol for calculating the Euclidean Distance of l points in Section 4.

2. THE PROPOSED SECURE PROTOCOLS FOR PRIVATE TRAJECTORY INTERSECTION TESTING

In the proposed secure protocols for two moving objects, the time-dependent polynomial function of the trajectory curve of each moving object should be first computed. Intersection takes place if the distance between two objects at a specific time point, t, is smaller than or equal to d. This constraint is applied to the whole time period and trajectory. Therefore, the trajectory function should also include other points in its proximity (with a radius d) at a specific time, t. Each party considers a point P(x(t), y(t)) on its corresponding curve at time t, preferring that no other parties are present in the proximity (with a radius d) of the point P at a specific time t, as illustrated in Figure 1.

To establish such conditions, one can represent the sur-rounding region (with a radius *d*) of a point at time *t*, as a circle in the two-dimensional space. Time is a key parameter in the proposed protocols, and intersections are considered to occur within the proximity region

Figure 1. The zone of non-intersection of moving objects

(with a radius d) of any point of the curve at a specific time, t. $P(x(t), y(t))$ is the center of circle (as shown in Figure 1) and hence, no other party (i.e. no other point) should be present in this region.

If the function of trajectory be $\vec{f}(t) = x(t).\hat{i} + y(t).\hat{j}$, we express the curve of this trajectory as $(x - x(t))^2 + (y - y(t))^2 = d^2$, which is based on limitation of predetermined threshold d.

In the next section, we present two protocols for PTIT, based on Gröbner Basis and Distance Computation.

2. 1. The Proposed Secure Protocol I- Based on Gröbner Basis

We want to obtain the intersection of the moving objects' curves, where the curves should be kept private. As the privacy of moving objects' trajectories is important, we generate new curves, where their intersections are the same as the intersections of the moving objects' curves. Moreover, by knowing the new generated curves, no information is obtained from the original curves except their intersections.

Hence, the privacy of the moving objects' trajectories is preserved. Also, in our proposed protocol, the dimension of space is not important.

In this section, we present an overview of a two-party PTIT (PTIT_Gröbner Protocol), which is based on Gröbner Basis. The details of this protocol was presented in Farokhi et al. [13].

PTIT_ GrÖbner Protocol

Protocol 1- The PTIT Based on GrÖbner Basis

party$_1$: selects $h_1(t)$ as random, computes $f_1(t)\times h_1(t)$

1) *Party$_1$ → party$_2$* : $f_1(t)\times h_1(t)$

party$_2$: selects $h_2(t)$ as random, computes $f_2(t)\times h_2(t)$

2) *Party$_2$ → party$_1$* : $f_2(t)\times h_2(t)$

party$_1$: calculates $V(f_1(t)\times h_1(t)\ ,\ f_2(t)\times h_2(t))$, where $V(f_1(t)\times h_1(t)\ ,\ f_2(t)\times h_2(t)) = ((a_1, a_2, \ldots, a_n), (a'_1, a'_2, \ldots, a'_n), \ldots, (a^{(n)}_1, a^{(n)}_2, \ldots, a^{(n)}_n))$

calculates $V(f_1(t) \times h_1(t))$ and removes them from $((a_1, a_2, \ldots, a_n), (a'_1, a'_2, \ldots, a'_n), \ldots, (a^{(n)}_1, a^{(n)}_2, \ldots, a^{(n)}_n))$ where the result is V_1

party$_2$: calculates $V(f_1(t)\times h_1(t)\ ,\ f_2(t)\times h_2(t))$, where $V(f_1(t)\times h_1(t)\ ,\ f_2(t)\times h_2(t)) = ((a_1, a_2, \ldots, a_n), (a'_1, a'_2, \ldots, a'_n), \ldots, (a^{(n)}_1, a^{(n)}_2, \ldots, a^{(n)}_n))$

calculates $V(,\ f_2(t)\times h_2(t))$ and removes them from $((a_1, a_2, \ldots, a_n), (a'_1, a'_2, \ldots, a'_n), \ldots, (a^{(n)}_1, a^{(n)}_2, \ldots, a^{(n)}_n))$ where the result is V_2

3) *Party$_1$ $\overset{SSI}{\leftrightarrow}$ party$_2$: Secure Set Intersection Protocol on V_1 and V_2*

2. 2. The Proposed Secure Protocol II- Based on Distance Computation of Curves

In this protocol, the parties first compute securely and jointly the distance function of their trajectories curves as $d(t)$. Each party can check if $d(t) \leq d$, by replacing the corresponding time in distance function $d(t)$. Thus, each party can obtain the collision points by itself.

As it is shown in Protocol 2, the distance function $d(t)$ includes three sections. Party$_1$ computes the first section of distance function which is $(x_1(t)^2 + y_1(t)^2)$. Party$_2$ computes the second section of distance function which is $(x_2(t)^2 + y_2(t)^2)$, and the last section should be computed jointly between parties which is $(x_1(t) \times x_2(t) + y_1(t) \times y_2(t))$.

In the first step of protocol, party$_1$ wishes to conceal $(x_1(t)^2 + y_1(t)^2)$. So, it composes an array containing the square of arbitrary values as M, where the k^{th} member of M is $(x_1(t)^2 + y_1(t)^2)$. Also, it adds the values of M by a random array as R. Then, M+R is sent to party$_2$.

In the second step of protocol, party$_1$ hides $(x_1(t), y_1(t))$ using two arrays A, B, so that the i^{th} member of A is $x_1(t)$ and the i^{th} member of B is $y_1(t)$. It should be noted that $(x_1(t), y_1(t))$ is used in the last section of distance function. Afterward, party$_1$ sends A, B to party$_2$.

Party$_2$ computes the multiplication of $(-2 \times x_2(t))$ by A, and $(-2 \times y_2(t))$ by B. It also adds the results together, and obtains $D = (-2 \times x_2(t) \times A) + (-2 \times y_2(t) \times B)$, that is the last part of distance function. Party$_2$ adds the values of C by a random noise $r_1(t)$.

In the third step of protocol, party$_1$ needs the i^{th} member of array C. Therefore, it requests party$_2$ using oblivious transfer. Therefore, party$_2$ sends the i^{th} member of C, without disclosure of its index and content.

In the fourth step of the protocol, party$_2$ adds $(x_2(t)^2 + y_2(t)^2) + r_2(t)$ to each member of the array sent by party$_1$ in the first step of protocol and obtains $[(M + R) + (x_2(t)^2 + y_2(t)^2) + r_2(t))\ ;\ 1 \leq i \leq n]$.

As $r_2(t)$ is a random noise function, it can be omitted by filtering. Among these results, party$_1$ just needs the k^{th} member of $[(M + R) + (x_2(t)^2 + y_2(t)^2) + r_2(t))\ ;\ 1 \leq i \leq n]$. So, it requests party$_2$ using oblivious transfer. Party$_2$ sends $(x_1(t)^2 + y_1(t)^2 + r^{(k)}(t)) + (x_2(t)^2 + y_2(t)^2 + r_2(t))$ without disclosure of its index and content.

Finally, party$_1$ computes the sum of $(x_1(t)^2 + y_1(t)^2 + r^{(k)}(t)) + (x_2(t)^2 + y_2(t)^2 + r_2(t))$-$(2 \times x_1(t) \times x_2(t) + 2 \times y_1(t) \times y_2(t) + r_1(t))$, and omits noises $r_1(t)$, $r_2(t)$, $r^{(k)}(t)$ by filtering. Thus, the distance function is computed securely and jointly.

It should be noted that the noises $r_1(t)$, $r_2(t)$ and array R are random signals such as white noises, with zero mean and finite variance. They have equal intensity and high fluctuation at different frequencies.

To know more about the Oblivious Transfer Protocol and its extension, an interested reader is referred to [27–30].

PTIT_Distances Protocol

Protocol 2- The PTIT Based on Distance Computation

$$d(t)=\sqrt{(x_1(t)-x_2(t))^2+(y_1(t)-y_2(t))^2}=\sqrt{x_1(t)^2+x_2(t)^2-2x_1(t)x_2(t)+y_1(t)^2+y_2(t)^2}$$
$$=\sqrt{-2y_1(t)y_2(t)}=(\underbrace{x_1(t)^2+y_1(t)^2}_{1})+(\underbrace{x_2(t)^2+y_2(t)^2}_{2})-2\underbrace{(x_1(t)x_2(t)+y_1(t)y_2(t))}_{3}$$

1) $party_1 \rightarrow party_2$: **M+R** *where* $\mathbf{M} = (a_1(t)^2, a_2(t)^2, \ldots, a_n(t)^2)$; **R** *is an array of random values, we denote them as* $(r^{(1)}(t), r^{(2)}(t), \ldots, r^{(n)}(t))$

$a_k(t)^2 = x_1(t)^2 + y_1(t)^2$; $1 \le k \le n$

2) $Party_1 \rightarrow party_2$: $\mathbf{A} = (x'_1(t), x_1''(t), \ldots, x_1^{(n)}(t))$;

$x_1^{(i)}(t) = x_1(t)$

$\mathbf{B} = (y'_1(t), y_1''(t), \ldots, y_1^{(n)}(t))$; $y_1^{(i)}(t) = y_1(t)$; $1 \le i \le n$

$party_2$: $\mathbf{C} = -2 \times (x_2(t) \times \mathbf{A}) - 2 \times (y_2(t) \times \mathbf{B})$

3) $Party_2 \overset{OT}{\leftrightarrow} party_1$: $-2 \times x_1(t) \times x_2(t) - 2 \times y_1(t) \times y_2(t) + r_1(t)$

$party_2$: $(\mathbf{M}+\mathbf{R}) + (x_2(t)^2 + y_2(t)^2 + r_2(t))$; $r_1(t)$, $r_2(t)$ *are random noises*

4) $party_2 \overset{OT}{\leftrightarrow} party_1$: $(a_k(t)^2 + r^{(k)}(t)) + (x_2(t)^2 + y_2(t)^2 + r_2(t)) = (x_1(t)^2 + y_1(t)^2 + r^{(k)}(t)) + (x_2(t)^2 + y_2(t)^2 + r_2(t))$

$party_1$: removes noises by filtering and computes $d^2(t) = (x_1(t)^2 + x_2(t)^2 - 2 \times x_1(t) \times x_2(t) + y_1(t)^2 + y_2(t)^2 - 2 \times y_1(t) \times y_2(t))$

3. SECURITY PROOF

In this section, we demonstrate the security proof of Protocol 2, based on the ideal-real simulation paradigm. The security proof of Protocol 1 was demonstrated in in [24].

3. 1. Security Proof of PTIT_Distance Protocol

We show that PTIT_Distance Protocol is secure, as long as the employed random generation system and oblivious transfer protocol is secure. Moreover, if the employed random generation system and oblivious transfer protocol are secure against semi-honest adversaries, our PTIT_Distance Protocol is also secure against semi-honest adversaries.

The following Theorem formalizes this statement.

Theorem 1. The PTIT_Distance Protocol is fully secure against a semi-honest adversary in the OT-hybrid model.

Proof. Our proof follows the ideal/real world paradigm [31]. In particular, we describe a simulator S_i that simulates the i^{th} party's view in the ideal world. Without loss of generality, we describe our security proof for the condition where simulator S_1 simulates the view of $party_1$ as Equation (1).

$$(f_1(t), g_1(t), g_2(t)), \tag{1}$$

where $f_1(t)$ is the input of $party_1$ i.e. $(x_1(t), y_1(t))$, and $g_1(t)$ is a random function, which is the output of oblivious transfer protocol in Step 3 instead of $-2 \times x_1(t) \times x_2(t) - 2 \times y_1(t) \times y_2(t) + r_1(t)$. In addition, $g_2(t)$ is a random function, which is the output of oblivious transfer protocol in Step 4 instead of $(x_1(t)^2+y_1(t)^2+r(k)(t))+(x_2(t)^2+ y_2(t)^2+ r_2(t))$.

We now prove Equation (2), which states that the view of party1 can be simulated by a probabilistic polynomial-time algorithm given access to the party1's input and output only. We emphasize that the adversary here is semi-honest, and therefore the view is exactly according to the protocol definition.

$$\{S_1(x, f(x,y))\}_{x,y\in\{0,1\}^*} \overset{c}{\equiv} \{view_1^{\pi}(x,y)\}_{x,y\in\{0,1\}^*} \tag{2}$$

Indistinguishability of views. We now prove by contradiction. Assume that there exists a probabilistic polynomial-time distinguisher D and a polynomial p(.) such that, for given *n*,

$$|\Pr[D(H_0(f_1(t), f_2(t), g_1(t), g_2(t)))=1]-\Pr[D(H_n(f_1(t), f_2(t), g_1(t), g_2(t)))=1]| > \frac{1}{p(n)} \tag{3}$$

where $f_1(t)$, $f_2(t)$, $g_1(t)$, $g_2(t)$ are given. $f_1(t)$, $f_2(t)$ are the inputs of $party_1$ and $party_2$, respectively. Also, $g_1(t)$, $g_2(t)$ are instead of the transferred messages.

We now use D to contradict the security of Step 2 of our proposed protocol. First, consider that the only difference between H_i and H_{i+1} is that the $(i + 1)^{th}$ run of this step in H_i is based on $(-2\times x_1(t)\times x_2(t)-2\times y_1(t)\times y_2(t),(x_1(t)^2+y_1(t)^2+r^{(k)}(t))+(x_2(t)^2+y_2(t)^2+r_2(t)))$ and the $(i + 1)^{th}$ run of this step in H_{i+1} is based on $(f_1(t), g_1(t), g_2(t))$, where $g_1(t)$, $g_2(t)$ are random values. Furthermore, given $(f_1(t), f_2(t), g_1(t), g_2(t))$ and a view *v*, it is possible to construct a distribution H such that if *v* is from $(-2\times x_1(t)\times x_2(t)-2\times y_1(t)\times y_2(t),(x_1(t)^2+y_1(t)^2+r^{(k)}(t))+(x_2(t)^2+y_2(t)^2+r_2(t)))$ then H = H_i, and if *v* is from $(f_1(t), g_1(t), g_2(t))$, then H = H_{i+1}. It therefore follows that it is possible to distinguish the view of $party_1$ in a real execution of our proposed protocol from its simulated view with the same probability that it is possible to distinguish H_i from H_{i+1}. However, this contradicts our assumption that $g_1(t)$ and $g_2(t)$ are randomly generated and contradicts the security of the random generation of polynomial functions and oblivious transfer protocol.

4. COMPLEXITY ANALYSIS

In this section, we analyze and compare the complexity of PTIT_GrÖbner Protocol and Garbled Circuit-based Protocol for Euclidean Distance Computation of l points.

4. 1. Complexity Analysis of PTIT_Gröbner Protocol

In order to investigate the computational complexity for the PTIT_Gröbner Protocol, we express the calculations in the protocol based on the binary operations such as 'and', 'or' and ‘xor’. In addition, to investigate the communication complexity, we calculate the number of transfer bits in each step of PTIT_ Gröbner Protocol. In the following, further details are expressed to analyze the complexity of the proposed protocol for PTIT_Gröbner Protocol. The steps outlined below are based on Protocol 1.

Step 1- $Party_1$ sends $f_1(t) \times h_1(t)$ to $party_2$. To transfer these arrays, it is necessary to send the coefficients of these polynomials in an L member array. Here, L is a security parameter and both parties are aware of its size. Since it is possible to have floating point coefficients in polynomials, each party can display the coefficients with a good accuracy using 32 bits. We consider the number of bits of the coefficients is equal to *m*. Thus, the transfer array is a (m × L)-bit array. It is also necessary to multiply two polynomials, which is multiplication of two (m × L)-bit arrays. This multiplication requires performing m × m × L × L times ‘and’ operations and m × L times ‘or’ operations, and should be done for each member of the array.

Step 2 - Similarly, $party_2$ sends $f_2(t) \times h_2(t)$ to $party_1$. Similarly, the transfer array is a (m × L)-bit array. It is also necessary to multiply two polynomials, which is multiplication of two (m×L)-bit arrays. This multiplication requires performing m × m × L × L ‘and’ operations and m × L ‘or’ operations.

Step 3 - Then, each party calculates the varieties of $h_2(t) \times f_2(t)$ and $h_1(t) \times f_1(t)$. The varieties are calculated using the Buchberger’s algorithm [32], which consists of a number of polynomial divisions. The polynomial division consists of binary multiplication and addition, which is reduced to modulo *n* at each step. Furthermore, since the coefficients of the polynomials have *m* bits, the division operation is performed in $G(m^n)$.

In each division operation, it is necessary to perform the basic ‘and / or’ operations for a number of times (which is equal to the bits of coefficients). Since the coefficients are *m* bits, it is necessary to perform ‘and’ , ‘or’ operations, m^2 and *m* times, respectively. In this manner, the Gröbner Basis of Ideal $< f_1(t) \times h_1(t), f_2(t), h_2(t) >$ is calculated. To find all bases, it is necessary to perform the division operation multiple times, equal to the number of ideal bases, which is assumed to be *r*. In other words, we should perform ‘and/ or’ operations L × (m^2 + m) × r times. This is carried out by each of the parties.

Step 4 - The first party eliminates the roots of $h_1(t)$ from the varieties of $< f_1(t) \times h_1(t), f_2(t), h_2(t) >$, and keeps the result secret. In fact, the first party acquires the common roots of $f_1(t)$ and $f_2(t) \times h_2(t)$. In this step and the fifth step, if we consider the number of ideal bases to be equal to s, L × (m^2 + m) × s basic ‘and’ or ‘or’ operations should be performed. The third, fourth, and fifth steps are performed without communication complexity, being only performed with computation complexity.

Step 5 - The second party removes the roots of $h_2(t)$ from the varieties of $< f_1(t) \times h_1(t), f_2(t), h_2(t) >$, and keeps the result secret, as well. Indeed, the second party obtains the common roots of $f_2(t)$ and $f_1(t) \times h_1(t)$. As stated in the fourth step, in this step, L × (m^2 + m) × s basic ‘and’ or ‘or’ operations should be performed.

Step 6 - Finally, both parties participate in the secure set intersection protocol, and securely obtain the common roots of $f_1(t)$ and $f_2(t)$. This step is done in combination, using a trusted third party. To implement this hybrid protocol, each party maintains its status in the main protocol as σ_i, and then executes the ρ_i protocol that securely calculates the intersection of sets. After receiving the output of ρ_i, they return to σ_i. To express the computation complexity of ρ_i, we use the secure protocol proposed in the literature [33]. In that protocol, each party forms a polynomial from the members of its own set, so that the members of the set are the roots of polynomial $(f_1(x) = (x - \alpha_1)(x - \alpha_2)...(x - \alpha_{k_1}))$. If the number of the set members is equal to k1, then the degree of $f_1(x)$ is also equal to k_1. Then, each party selects a random polynomial r(x). Each party shares the coefficients of the polynomial h(x)=f(x).r(x) to the other party. Each party can use the share values to create a part of the polynomial $u(x)=h_1(x) + h_2(x)$ and send it to the other party. With calculating the final polynomials, the parties can substitute the members of their own sets into the polynomial. If the evaluation is zero then the element belongs to the intersection, else the element does not belong to the intersection. These calculations are a number of addition and multiplication operations that can be converted to ‘and /or’ operations. The numbers of addition and multiplication operations is in order of $O(k_1 + k_2)$, where k_1 is the number of set members of the first party, and k_2 is the number of set members and the polynomial degree of the second party, respectively. If we represent each coefficient with m bits, each addition operation can be converted to m times ‘or’ operations. Moreover, each multiplication operation can be converted to m times ‘or’ operations and m^2 times ‘and’ operations. Finally, the computation complexity of the PTIT_Gröbner Protocol is in order of $O(m^2 \times (k_1 + k_2))$. In the worst case, the size of the two sets are equal, which can be considered to be k. In this case, the computation complexity is in order of $O(m^2 \times k)$. To calculate the number of transfer bits between parties, we need to calculate the number of bits to share the coefficients of polynomials. Each party creates the coefficient shares and then, sends it to the other party. This action is done twice by each party and hence, a total of 2 × (k_1 + k_2) × m bits are sent. In general, the communication complexity of secure set intersection protocol is in order of O(k × m). Furthermore, each party enters its own set into ρ_i protocol such that the size of inputs is 2 × k × m bits at worst.

Finally, the computation complexity of PTIT_Gröbner Protocol is as follows, which indicates the number of 'and , or' operations: $2 \times (m^2 + m) \times L + 2 \times (m^2 + m) \times r \times L + 2 \times (m^2 + m) \times s \times L + m^2 \times k \approx O(m^2 \times L \times (r + s + k)) \approx O(m^2 \times L^2 \times Q)$.

In the above computation complexity, L indicates the array size of the polynomial coefficients, r and s denote the number of the bases generating the ideals of polynomials in the third to fifth steps, and k represents the number of varieties that enter the secure set intersection protocol. We denote the summation of these variables with Q.

If the degree of polynomials is finite, the number of polynomial coefficients, which is denoted as L, has also a finite size. On the other hand, since the degree of ideals and the number of variables are finite, the number of bases of ideals in the third to fifth steps is finite, too. In addition, the number of input bits for displaying the location coordinates is finite as well, such that decimal numbers can be displayed with a good accuracy using 32 bits. If the degree of polynomials are assumed to be finite, the computation complexity of this protocol is of numerical order.

The communication complexity of the protocol steps is as $2 \times m \times L + 3 \times k \times m \approx O((k + L) \times m)$ transfer bits, in which L is the size of the coefficients array. Also, k is the number of varieties, which are the inputs of the secure set intersection protocol, and m is the number of the coefficients bits, which are used in the secure set intersection protocol. If we assume that the degree of the polynomials is finite, L and k are finite. Moreover, if we consider the number of bits used for displaying the coefficients to be 32 bits, the communication complexity of this protocol is numerical.

4. 2. Complexity Analysis of Garbled Circuit-based Protocol for Euclidean Distance Computation of l Points One way to analyze the complexity of secure multi-party computation protocols is to compare the complexity of the proposed protocols with those of the Yao's garbled circuit-based protocols [2].

The basic operation for finding the intersection of two moving objects' trajectories regardless of security requirement can be considered in such a way that each moving object has l discrete points.

In this regard, basic operations are implemented using the Yao's garbled circuit [25, 26] (regardless of the proposed security approach) and its computation and communication complexities are examined and compared with the secure proposed protocol.

The details of Yao's garbled circuit have been previously presented [25, 26]. In this section, we express the computation and communication complexities of this circuit to calculate the Euclidean Distance of *l* points.

To create a garbled circuit for calculating the Euclidean Distance, it is necessary to consider the number of input bits and the number of base operations for computing the Euclidean Distance. The number of bits for representation of the points P and Q are considered to be equal to m. Since x and y coordinates are used to represent spatial points (P and Q), it is necessary to display each coordinate with m bits and calculate the Euclidean distance as Equation (4).

$$d(t)=\sqrt{(x_1(t)-x_2(t))^2+(y_1(t)-y_2(t))^2} = \{x_1(t)^2+x_2(t)^2+y_1(t)^2+y_2(t)^2 - 2\times(x_1(t)x_2(t)+y_1(t)y_2(t))\}^{\frac{1}{2}} \tag{4}$$

Each party performs the square operation on its side. Then, (m+1)-bit addition carried out five times, and m-bit multiplication are done two times, where m is the number of input bits. The (m + 1)-bit addition includes m + 1 times 'or' operations, and the m-bit multiplication includes m^2 times 'and', and also m times 'or' operations.

The sender creates a circuit with a number of gates described. It then creates a truth table for each gate and garbles it. For garbling, it is necessary to create independent keys for the inputs and outputs of 'and' operations ($2m^2$ times) and 'or' operations (5m + 5 times). In fact, it creates $6 \times (5m + 5 + 2m^2) \approx 12m^2$ keys, which are n-bit keys. Since these gates include 'and / or' gates, it is necessary to create keys and garbled truth tables for all gates.

For each gate, the sender creates a garbled truth table that includes output keys encrypted using the corresponding input keys. According to the Half Gates method [34] that uses Row Reduction [35], for each gate, it is necessary to perform the modular exponentiation only twice, and send two n-bit messages that are the outputs of the modular exponentiation. Indeed, $2 \times (5m + 5 + 2m^2) \times (n^6 + n^3 + n)$ 'and' and 'or' operations are required to create a garbled truth table. Furthermore, $2 \times (5m + 5 + 2m^2) \times n$ bits are transferred.

For each gate, the sender sends one n-bit key corresponding to its input to the receiver. The number of these gates is $(5m + 5 + 2m^2)$ and for each gate, one n-bit key is transferred. The number of transfer bits equals to $(5m + 5 + 2m^2) \times n$.

In addition, in the oblivious transfer protocol, the sender and receiver transfer a finite number of bits, and the final output of the oblivious transfer protocol is an n-bit key. For each 'and / or' gate, an oblivious transfer protocol is performed, and one n-bit key is transferred. Therefore, a total of $(5m + 5 + 2m^2) \times n$ bits are transferred. Thus, the communication complexity of the protocol for finding the Euclidean Distance between two m-bit points using garbled circuit equals: 4 × n ×

$(5m + 5 + 2m^2) = 20 \times n \times m + 20 \times n + 8 \times m^2 \times n \approx O(m^2 \times n)$ bits.

To calculate the computation complexity, we consider the calculations that the sender and receiver do on their side. The sender generates the C circuit, which describes the Euclidean Distance function. It includes $6m + 4$ ‘or’ gates and $2m^2$ ‘and’ gates. It then creates n-bit keys for each input and keeps them secret. As previously described, a modular exponentiation algorithm such as RSA is used to create the n-bit keys. So, it is required to perform ‘and’ operations n^6 times, and ‘or’ operations $n^3 + n$ times.

The sender generates six keys for each ‘and / or’ gates, where four keys are generated for inputs and two keys are for output. Totally, the sender generates $6 \times (6m + 4 + 2m^2)$ n-bit keys. Creating such keys requires $6 \times (6m + 4 + 2m^2) \times (n^3 + n + n^6) \approx O(n^6 \times m^2)$ ‘and / or’ operations.

For each gate, the sender creates a garbled truth table that includes output keys encrypted using the corresponding input keys. He uses a modular exponentiation algorithm such as RSA to encrypt the output keys. We assume that the input and output keys and all security parameters have n bits. The sender can use the fast modular exponentiation algorithm for encrypting, which requires n^6 ‘and’ operations and $n^3 + n$ ‘or’ operations for creating each member of the garbled truth table. The number of key bits generated by the sender equals $O(n^6 \times m^2) \approx (n^6 + n^3 + n) \times 6 \times (6m + 2m^2)$ bits.

In each implementation of the oblivious transfer protocol, the sender and receiver perform three ‘xor’ operations on their own sides. We apply Free Xor [36, 37] method to reduce the communication cost per each ‘xor’ gate. In addition, in Free Xor method the number of generated keys per ‘xor’ gate, the number of transferred keys, and the number of encryption and decryption are zero. In fact, the garbled ‘xor’ operation is converted to a simple ‘xor’.

In addition, the permutation is executed three times. If we consider the permutation as the modular exponentiation, and use the fast modular exponentiation algorithm, ‘and’ operation should be done n^6 times, and ‘or’ operations should be done $n^3 + n$ times. Thus, for each gate, an oblivious transfer protocol is performed, where $(n^6 + n^3 + n) \times (6m + 2m^2) \approx O(n^6 \times m^2)$ times ‘and / or’ operations are done for it.

In addition, the sender and receiver perform the x and y square operations, including multiplication and addition operations. If this operation is performed in a binary manner, for each m-bit multiplication operation, m^2 times ‘and’, m times ‘or’ operations are performed. On the other hand, for each m-bit addition operation, m times ‘or’ operations are performed. Therefore, a total of $2 \times (m^2 + 2m) \approx m^2$ ‘and/or’ operations would be generally needed.

Finally, the computation complexity of Yao's garbled circuit for calculating the Euclidean Distance of two points is in order of $n^6 \times m^2$. The computation and communication complexity that we have expressed is for calculating the Euclidean Distance of two m-bit points. On the other hand, the trajectory of moving objects includes l points, which are m-bit. It is necessary to perform the above operation l times. Therefore, the computation complexity is in order of $n^6 \times m^2 \times l$, and the communication complexity is in order of $m^2 \times n \times l$, depending on the number of input bits, key size and the number of trajectory points. Increasing these three parameters will increase the computation and communication complexities.

4. 3. Complexity Comparison of PTIT_Gröbner Protocol and Garbled Circuit_based Protocol
The input bits represent the x and y coordinates of the trajectory points. These points can be represented by decimal numbers, which can be accurately indicated with 32 bits.

We can consider all variables (except for the security parameter, n) to be equal, due to the limitation and small growth compared to n, and calculate the complexities. In this manner, the computation complexity of the garbled circuit- based protocol is in order of n^6, and the order of its communication complexity is n. In contrast, the computation and communication complexities of the PTIT_Gröbner protocol is numeric.

As illustrated in Table 1, the PTIT_ Gröbner protocol is better than the garbled circuit-based protocol. Moreover, the PTIT_Distance Protocol is based on oblivious transfer protocol. Therefore, it is based on modular exponentiation, and we need ‘and’ operations n^6 times, and ‘or’ operations $n^3 + n$ times. Thus, due to high growth of n, its complexity is almost the same as the complexity of garbled circuit-based protocol.

As stated before, the problem of PPT is similar to PTIT. However, in PPT the parties should have smart devices with LBS. Thus, to compare the complexity of our proposed protocols and the previous decentralized protocols for PPT [11, 18–22], we demonstrate their complexities in Table 2.

All the mentioned protocols in Table 2 are based on modular exponentiation, and they need ‘and’ operations n^6 times, and ‘or’ operations $n^3 + n$ times, for each modular exponentiation. A comparison between Table 1 and Table 2 shows that PTIT_Gröbner protocol is better than the garbled circuit-based protocol for Euclidean Distance Computation of l points and the previous researches for PPT.

TABLE 1. Complexity comparison of our proposed protocols and garbled circuit-based protocol for Euclidean Distance calculation

Communication Complexity	Computation Complexity	Protocol
$O(m^2 \times n \times l)$	$O(n^6 \times m^2 \times l)$	Calculation of the Euclidean Distance of *l* points based on the Yao's garbled circuit
$O(m \times (k+L))$	$O(m^2 \times L \times Q)$	PTIT Protocol based on GrÖbner Basis (PTIT_ GrÖbner Protocol)
$O(N \times m)$	$O(n^6 \times m \times N)$	PTIT Protocol based on Distance Computation (PTIT_Distance Protocol)

n: The number of bits for input-output keys and security parameters

m: The number of input bits

Q: The sum of s + r + k, related to three variables, representing the number of varieties and the number of ideals at different stages.

k: The number of input varieties of the secure set intersection protocol (also, the degree of polynomials in the secure set intersection protocol)

L: The size of the polynomial coefficients array

N: The size of transferred array for concealing the trajectory points

TABLE 2. Complexity comparison of the proposed protocols for PPT [11, 13-15, 17, 22]

Number of Base Operations	Number of Exponentiation Operations	Cryptography Algorithm	Protocol
For each exponentiation, n^6 '*and*' operation, n^3+n '*or*' operation is required. Alice: $6\times(n^6+n^3+n)$+ 3 DL Bob: $6\times(n^6+n^3+n)$	Alice: 6 exp+3 DL Bob: 6 exp	CGS	Pierre, 2007 [18]
$3\times(n^6+n^3+n)$	3 exp/user	DH-based SPEKE	NFP, 2009 [20]
Alice: $3\times(n^6+n^3+n)$ Bob: $4\times(n^6+n^3+n)$	Alice: 3 exp Bob: 4 exp	ElGamal encryption	EG-PET , 2011 [19]
$3\times(n^6+n^3+n)$	3 exp/user	DH-based SPEKE	DH-PET, 2014 [21]
$2\times(n^6+n^3+n)$	2 exp/user	RSA	FP-PET , 2016 [22]
Alice: $3\times(n^6+n^3+n)$ Bob: $2\times(n^6+n^3+n)$	Alice: 3 exp Bob: 2 exp	DGK	FPPLP, 2018 [11]

5. CONCLUSION

In this paper, we presented two protocols for private intersection detection of two moving objects' trajectories, which are based on Gröbner Basis and Distance Computation. The Gröbner-based protocol was proposed in a previous research, and we only presented its overview. We also presented the complexity analysis of the Gröbner-based protocol. We compared its complexity by the garbled circuit-based protocol for Euclidean Distance Computation of *l* points. This approach for complexity analysis is common in the field of security protocol design. Moreover, we proved the security of the proposed protocol, which is based on Distance Computation.

Our comparison demonstrated that PTIT_Gröbner protocol is better than the garbled circuit-based protocol for Euclidean Distance Computation of l points and the previous researches for PPT.

6. REFERENCES

1. Asadi Saeed Abad, F., and Hamidi, H., "An Architecture for Security and Protection of Big Data", ***International Journal of Engineering - Transaction A: Basics***, Vol. 30, No. 10, (2017), 1479–1486. doi:10.5829/ije.2017.30.10a.08
2. Yao, A. C., "Protocols for secure computations", 23rd Annual Symposium on Foundations of Computer Science (Sfcs 1982), (1982), IEEE, 160–164. doi:10.1109/SFCS.1982.38
3. Goldreich, O., Micali, S., and Wigderson, A., "How to play ANY mental game", Proceedings of the Nineteenth Annual ACM Conference on Theory of Computing - STOC '87, (1987), 218–229. doi:10.1145/28395.28420
4. Hemenway, B., Lu, S., Ostrovsky, R., and Welser IV, W., "High-Precision Secure Computation of Satellite Collision Probabilities", In International Conference on Security and Cryptography for Networks, Springer, Cham, (2016), 169–187. doi:10.1007/978-3-319-44618-9_9
5. Atallah, M. J., and Du, W., "Secure Multi-party Computational Geometry", In Workshop on Algorithms and Data Structures, Springer, Berlin, Heidelberg, (2001), 165–179. doi:10.1007/3-540-44634-6_16

6. Frikken, K. B., and Atallah, M. J., "Privacy preserving route planning", Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society - WPES ’04, (2004). doi:10.1145/1029179.1029182
7. Cheng, R., Zhang, Y., Bertino, E., and Prabhakar, S., "Preserving User Location Privacy in Mobile Data Management Infrastructures", In International Workshop on Privacy Enhancing Technologies, Springer, Berlin, Heidelberg, (2006), 393–412. doi:10.1007/11957454_23
8. Mascetti, S., Freni, D., Bettini, C., Wang, X. S., and Jajodia, S., "Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies", ***The VLDB Journal***, Vol. 20, No. 4, (2011), 541–566. doi:10.1007/s00778-010-0213-7
9. Zhu, H., Wang, F., Lu, R., Liu, F., Fu, G., and Li, H., "Efficient and Privacy-Preserving Proximity Detection Schemes for Social Applications", ***IEEE Internet of Things Journal***, Vol. 5, No. 4, (2018), 2947–2957. doi:10.1109/JIOT.2017.2766701
10. Zheng, Y., Li, M., Lou, W., and Hou, Y. T., "Location Based Handshake and Private Proximity Test with Location Tags", ***IEEE Transactions on Dependable and Secure Computing***, Vol. 14, No. 4, (2017), 406–419. doi:10.1109/TDSC.2015.2472529
11. Järvinen, K., Kiss, Á., Schneider, T., Tkachenko, O., and Yang, Z., "Faster Privacy-Preserving Location Proximity Schemes", In International Conference on Cryptology and Network Security,Springer, Cham, (2018), 3–22. doi:10.1007/978-3-030-00434-7_1
12. Pagnin, E., Gunnarsson, G., Talebi, P., Orlandi, C., and Sabelfeld, A., "TOPPool: Time-aware Optimized Privacy-Preserving Ridesharing", ***Proceedings on Privacy Enhancing Technologies***, Vol. 2019, No. 4, (2019), 93–111. doi:10.2478/popets-2019-0060
13. Farokhi, F., Shames, I., and Johansson, K. H., "Private routing and ride-sharing using homomorphic encryption", ***IET Cyber-Physical Systems: Theory & Applications***, Vol. 5, No. 4, (2020), 311–320. doi:10.1049/iet-cps.2019.0042
14. Oleynikov, I., Pagnin, E., and Sabelfeld, A., "Where Are You Bob? Privacy-Preserving Proximity Testing with a Napping Party", In European Symposium on Research in Computer Security, Springer, Cham, (2020), 677–697. doi:10.1007/978-3-030-58951-6_33
15. Chor, B., Kushilevitz, E., Goldreich, O., and Sudan, M., "Private information retrieval", ***Journal of the ACM***, Vol. 45, No. 6, (1998), 965–981. doi:10.1145/293347.293350
16. Siksnys, L., Thomsen, J. R., Šaltenis, S., and Yiu, M. L., "Private and Flexible Proximity Detection in Mobile Social Networks", 2010 Eleventh International Conference on Mobile Data Management, (2010), IEEE, 75–84. doi:10.1109/MDM.2010.43
17. Šikšnys, L., Thomsen, J. R., Šaltenis, S., Yiu, M. L., and Andersen, O., "A Location Privacy Aware Friend Locator", In International Symposium on Spatial and Temporal Databases, Springer, Berlin, Heidelberg., (2009), 405–410. doi:10.1007/978-3-642-02982-0_29
18. Zhong, G., Goldberg, I., and Hengartner, U., "Louis, Lester and Pierre: Three Protocols for Location Privacy", In International Workshop on Privacy Enhancing Technologies, Springer, Berlin, Heidelberg, (2007), 62–76. doi:10.1007/978-3-540-75551-7_5
19. Narayanan, A., Thiagarajan, N., Lakhani, M., Hamburg, M., and Boneh, D., "Location Privacy via Private Proximity Testing", In NDSS, Vol. 11, (2011).
20. Chatterjee, S., Karabina, K., and Menezes, A., "A New Protocol for the Nearby Friend Problem", In IMA International Conference on Cryptography and Coding, Springer, Berlin, Heidelberg, (2009), 236–251. doi:10.1007/978-3-642-10868-6_14
21. Magkos, E., Kotzanikolaou, P., Magioladitis, M., Sioutas, S., and Verykios, V. S., "Towards Secure and Practical Location Privacy through Private Equality Testing", In International Conference on Privacy in Statistical Databases, Springer, Cham, (2014), 312–325. doi:10.1007/978-3-319-11257-2_24
22. Kotzanikolaou, P., Patsakis, C., Magkos, E., and Korakakis, M., "Lightweight private proximity testing for geospatial social networks", ***Computer Communications***, Vol. 73, (2016), 263–270. doi:10.1016/j.comcom.2015.07.017
23. Šeděnka, J., and Gasti, P., "Privacy-preserving distance computation and proximity testing on earth, done right", Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, (2014), 99–110. doi:10.1145/2590296.2590307
24. Dehghan, M., and Sadeghiyan, B., "Privacy-preserving collision detection of moving objects", ***Transactions on Emerging Telecommunications Technologies***, Vol. 30, No. 3, (2019), e3484. doi:10.1002/ett.3484
25. Huang, Y., Evans, D., and Katz, J., "Private set intersection: Are garbled circuits better than custom protocols?", In 19th Network and Distributed Security Symposium, San Diego, (2012), 1–15.
26. Bellare, M., Hoang, V. T., and Rogaway, P., "Foundations of garbled circuits", Proceedings of the 2012 ACM Conference on Computer and Communications Security - CCS ’12, (2012), 784–796. doi:10.1145/2382196.2382279
27. Rabin, M. O., "How To Exchange Secrets with Oblivious Transfer", IACR Cryptol. EPrint Arch., (2005).
28. Yao, A. C.-C., "How to generate and exchange secrets", 27th Annual Symposium on Foundations of Computer Science (Sfcs 1986), (1986), IEEE, 162–167. doi:10.1109/SFCS.1986.25
29. Naor, M., and Pinkas, B., "Efficient oblivious transfer protocols", In Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms, Washington DC (Vol. 1), (2001), 448–457. doi:10.1145/365411.365502
30. Ishai, Y., Kilian, J., Nissim, K., and Petrank, E., "Extending Oblivious Transfers Efficiently", In Annual International Cryptology Conference, Springer, Berlin, Heidelberg, (2003), 145–161. doi:10.1007/978-3-540-45146-4_9
31. Lindell, Y., and Pinkas, B., "A Proof of Security of Yao’s Protocol for Two-Party Computation", ***Journal of Cryptology***, Vol. 22, No. 2, (2009), 161–188. doi:10.1007/s00145-008-9036-8
32. Cox, D. A., Little, J., and O’Shea, D., Ideals, Varieties, and Algorithms, Cham, Springer International Publishing, (2015). doi:10.1007/978-3-319-16721-3
33. Li, R., and Wu, C., "An Unconditionally Secure Protocol for Multi-Party Set Intersection", In International Conference on Applied Cryptography and Network Security, Springer, Berlin, Heidelberg, (2007), 226–236. doi:10.1007/978-3-540-72738-5_15
34. Zahur, S., Rosulek, M., and Evans, D., "Two Halves Make a Whole", In Annual International Conference on the Theory and Applications of Cryptographic Techniques, Springer, Berlin, Heidelberg, (2015), 220–250. doi:10.1007/978-3-662-46803-6_8
35. Pinkas, B., Schneider, T., Smart, N. P., and Williams, S. C., "Secure Two-Party Computation Is Practical", In International Conference on the Theory and Application of Cryptology and Information Security, Springer, Berlin, Heidelberg, (2009), 250–267. doi:10.1007/978-3-642-10366-7_15
36. Kolesnikov, V., and Schneider, T., "Improved Garbled Circuit: Free XOR Gates and Applications", In International Colloquium on Automata, Languages, and Programming, Springer, Berlin, Heidelberg., (2008), 486–498 Berlin, Heidelberg, Springer Berlin Heidelberg. doi:10.1007/978-3-540-70583-3_40
37. Kolesnikov, V., Mohassel, P., and Rosulek, M., "FleXOR: Flexible Garbling for XOR Gates That Beats Free-XOR", In Annual Cryptology Conference, Springer, Berlin, Heidelberg, (2014), 440–457. doi:10.1007/978-3-662-44381-1_25

Persian Abstract

چکیده

ما در این مقاله دو پروتکل برای یافتن امن اشتراک مسیر حرکت دو شیء متحرک ارائه می‌دهیم. برای طراحی پروتکل اول، مسئله یافتن امن اشتراک مسیر حرکت را به مسئله یافتن امن ریشه مشترک دو چند جمله‌ای نمایش دهنده مسیر حرکت تبدیل می‌نماییم. سپس، از پایه گروبنر برای طراحی پروتکل یافتن امن ریشه مشترک دو چند جمله ای استفاده می‌کنیم. همچنین، پروتکل دیگری بر مبنای محاسبه فاصله دو منحنی مسیر حرکت ارائه می‌دهیم. پروتکل یافتن امن ریشه مشترک دو چند جمله‌ای بر مبنای پایه گروبنر، دارای پیچیدگی عددی است و در مقایسه با پروتکل مبتنی بر مدار مبهم برای محاسبه فاصله اقلیدسی l نقطه و پروتکل های پیشین برای بررسی امن نزدیک بودن اشیاء متحرک، دارای پیچیدگی بسیار پایین‌تری است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Design and Analysis of High Efficiency Perovskite Solar Cells with Light Trapping Nano-textured Substrates

H. Abedini-Ahangarkola, S. Soleimani-Amiri*

Department of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran

PAPER INFO

Paper history:
Received 01 September 2020
Received in revised form 09 November 2020
Accepted 09 January 2021

Keywords:
Perovskite Solar Cells
Nano-textured
COMSOL Multiphysics
Finite Element Method

ABSTRACT

Recently, the utilization of hybrid organic-inorganic perovskite solar cells under advanced light management designs have attracted intensive attention. In this study, a three-dimensional (3D) finite element method (FEM) technique was used in the COMSOL Multiphysics simulation package to investigate coupled optical and electrical characteristics of perovskite solar cells (PSCs) with light trapping nanostructures. Upon the use of nano-textured fluorine-doped tin oxide (FTO) substrates, we propose two architectures which can guide and trap the light at nanometer dimensions. Two proposed PSCs i.e. concave and trapezoidal structures are compared to the planar structure in order to investigate the effects of using nanostructured substrates on the optoelectronic performance of PSCs. Optical analysis reveals that using optimized concave and trapezoidal structures can enhance the light absorption up to 32 and 26%, respectively at the wavelength of 550 nm. Electrical simulations have shown that in addition to enhanced total carrier generation, the generated carriers can be effectively collected in the proposed nanostructured PSCs. Accordingly, the short-circuit current has risen from 20 mA for planar structure to 25.7 mA for concave and 23.2 mA for trapezoidal PSCs. After analyzing various heights and adopting optimum values, the power conversion efficiency for concave and trapezoidal PSCs experienced substantial increase of 5.5 and 3.5%, compared to the planar structure. These drastic improvements analyzed by coupled optical and electrical modelling of nanostructures can pave the way for further studies to fabricate high efficiency PSCs with nano-textured substrates as a light-trapping technique.

***doi**: 10.5829/ije.2021.34.04a.13*

1. INTRODUCTION

Hybrid halide perovskite has attracted so much attention as a great photoactive material in photovoltaic device applications due to its large absorption coefficient [1, 2], an appropriate and tunable band gap [3], long carrier diffusion length and fabulous efficiency in harvesting energy [4]. Moreover, the simplicity of manufacturing process and the feasibility of being integrated with traditional photovoltaics such as Si and CIGS to build efficient tandem devices are some of the other advantages of perovskite solar cells (PSCs) which would surely outweigh their disadvantages, such as vulnerability due to exposure to moisture, heat, etc [5, 6]. In the course of the years 2009 to 2016, the efficiency of PSCs increased from 3.8 to above 22% [7-9]. The optimization of perovskite preparation and deposition methods, the enhancement of interface quality and degradation mitigation are the main approaches that have been used in the ongoing efforts to achieve PSCs with a greater efficiency. Besides, different device configurations were developed, including the original mesoporous n-i-p junction and mesoporous-free planar junction with regular (n-i-p) and inverted structure (p-i-n) [10, 11].

Although tremendous progress has been made, many challenges for PSCs still exist. The most important problem is the conflict between the amount of carrier photo-generation and collection. As we all know, the thickness of the perovskite layer should be less than the diffusion length of carriers for efficient extraction, whereas the thin perovskite material will lead to insufficient light absorption, which means fewer

*Corresponding Author Institutional Email: *s.soleimani@nit.ac.ir* (S. Soleimani Amiri)

generated carriers [12]. To solve the problem, thin absorbing layers in combination with light trapping techniques are exploited to increase light absorption within the active layer [13].

Generally, there are two typical techniques to enhance optical trapping capability. The first one is the light scattering by increasing the optical paths of the incident light, and the other one is to enhance the surface plasmon resonance effect [14, 15]. Efficient light trapping methods can also lead to decreased manufacturing time and cost. Therefore, acquiring more knowledge to realize light trapping effects in photo absorber layer is necessary for the future development of highly-efficient, environmentally friendly, and low-cost thin film solar cells. It is worth noting that the optical and electrical benefits of light trapping can vary for solar cells different materials. A wide variety of plasmonic and nanophotonic structures were proposed for the first and second generations of solar cells, and many theoretical analyses were carried out to model both electrical and optical behavior in these advanced photovoltaic devices to predict the performance and optimize the structure [16].

In the field of perovskite solar cell, however, studies are relatively limited about various light trapping techniques with experimentally realistic structures and their optical and electrical effects [17, 18]. Abdelraouf et al. [17] used nanotubes, nanorods, nanocones, and nanopyramids as hole transport nanostructures for light trapping and enhancing the PSC performance. They found that the nanotubular structures, which can be fabricated by free standing TiO_2 array, have the highest short circuit current (24.29 mA/cm^2) and overall efficiency (15.23%) among the different investigated nanostructures. Zandi et al. [18] explored the use of corrugated anti and back reflector layers instead of flat ones to improve the power conversion efficiency of PSCs, and demonstrated that in this way they can increase the PCE of their PSC to 17.5%.

In this context, we make an effort to improve the performance of $CH_3NH_3PbI_3$ PSCs using two different sets of nano-textured FTO substrates (periodic concave and trapezoidal structures) as a light trapping method. For each model, using Finite-Element-Method (FEM), we investigate the impact of nanostructures on the light trapping and carrier generation, and then couple the results with electrical model to calculate the short circuit current density (J_{sc}), open circuit voltage (V_{oc}), power conversion efficiency (PCE) and other important characteristics such as incident photon-to-electron conversion efficiency (IPCE). By changing the height, the optimum value of each nanostructure is chosen to demonstrate the superior function of the device and provide a guideline for further research on designing such delicate nano-textured substrates and make the most of light trapping effect.

2. MODELING DETAILS

Our base materials for this simulation work are Au, Spiro-MeOTAD, $CH_3NH_3PbI_3$, TiO_2, FTO and air. It is obvious that fluorine-doped tin oxide (FTO) acts as the transparent conducting oxide (TCO), TiO_2 as the electron transporter layer (ETL), Spiro-MeOTAD as the hole transport layer (HTL) and Au as contact.

The numerical optoelectronic simulations based on finite-element method (FEM) are carried out in COMSOL Multiphysics package. The FEM is a numerical method to calculate boundary condition problems for partial differential equations [18]. In this method, the whole structure is subdivided into small parts by mesh generation, and the equations are solved approximately for each element. For a combined optical-electrical simulation, first we use wave optic module to calculate optical absorption and generation profiles of electrons within the device. In fact, in this step, the Helmholtz equation derived from Maxwell's equation in the frequency domain are solved for each element to obtain optical intensity and photogeneration rate. The equation is given as:

$$\nabla \times (\nabla \times \mathrm{E}) - k_0^2\, \varepsilon_r\, E = 0$$
$$\varepsilon_r(\lambda) = (\, n(\lambda) \; - ik(\lambda))^2\,, \qquad (1)$$

where E is electric field, k is wave-vector and is relative permittivity ε_r which is a function of wavelength. Relative permittivity for each layer is calculated by real part (n) and imaginary part (k) of refractive index extracted from [19-21] and implemented in COMSOL Multiphysics using interpolation function. Standard AM1.5G plane wave source is located above the air and is used as the input power. The wavelength range in our simulation is 300-800 nm with 10 nm resolution. In order to reduce the simulation time, we model a small unit of solar cell with periodic boundary condition (PBC) on the sides of all PSC layers. Only for the Au contact, we define the layer as a perfect electric conductor (PEC).

Setting this condition, the gold acts as a back reflector and most of the unabsorbed photons of the incident light are reflected back to the structure. Using Equation (1), the electric field E can be obtained for the whole domain. Photo generation rate for each wavelength under AM1.5G irradiation can be calculated by (Equation (2)) [18]:

$$G_{opt}(\lambda) = \frac{\varepsilon'' |E|^2}{2\hbar}\,, \qquad (2)$$

where $\hbar$ is the plank constant and ε'' is the imaginary part of the permittivity. As stated before, complex refractive indices and relative permittivity of all layers are taken from previously measured data [18, 22]. By integration of $G_{opt}(\lambda)$ over the wavelength range, the total generation rate (G_{tot}) for the whole spectrum (300-800 nm) is calculated as follows:

$$G_{tot} = \int_{300}^{800} G_{opt}(\lambda)\ d\lambda \tag{3}$$

Then, the calculated total generation rate is used as an input of the semiconductor module to investigate electrical properties of the PSC. By solving Poisson (Equation (4)) and continuity equations (Equations (5), (6)) at each mesh element in semiconductor module the carrier concentration in the structure is obtained

$$\nabla^2\varphi = \frac{\rho}{\varepsilon_0\varepsilon_r} = \frac{q(p+N_D-n-N_A)}{\varepsilon_0\varepsilon_r} \tag{4}$$

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla\cdot J_n + G_n - R_n \tag{5}$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q}\nabla\cdot J_p + G_p - R_p \tag{6}$$

where φ is the electrostatic potential, ρ is the total charge density, ε_0 is the permittivity of free space, ε_r is the relative permittivity of semiconductor, q is the electric charge, p and n are the number of charge carriers, N_D and N_A are the concentrations of donor and acceptor impurities, G_n and G_p are the generation rates, and R_n and R_p are the recombination rates of electrons and holes, respectively. As stated before, generation rates $G_n = G_p = G_{tot}$ are imported from the wave optic module results (Equation (2)) to couple optical and electrical models. J_n and J_p are the electron and hole current density, expressed by the electron and hole drift-diffusion charge transport equations as follows:

$$J_n = qn\mu_n E + qD_n\nabla n \tag{7}$$

$$J_p = qp\mu_p E - qD_p\nabla p \tag{8}$$

where μ_n and μ_p are the electron and hole mobility, and D_n and D_p are the electron and hole diffusion coefficient, respectively.
The physical parameters of TiO_2, $CH_3NH_3PbI_3$ and Spiro-MeOTAD used in the semiconductor module are summarized in Table 1 [14, 18]. Shockley-Read-Hall (SRH) recombination, as the dominant recombination mechanism, is taken into consideration in our simulations. The FTO transparent contact and Au back reflector contact are assumed Ohmic ($10\frac{\Omega}{cm^2}$) and Schottky, respectively.

After obtaining the current density-voltage (J-V) characteristics of the solar cell, we can determine fill-factor (FF) and power conversion efficiency (PCE). The fill factor of the device is defined as:

$$FF = {(J_m \cdot V_m)}\Big/{(J_{sc} \cdot V_{oc})} \tag{9}$$

where J_m and V_m are the current density and voltage at the maximum power point (P_m), and J_{sc} is short-circuit current density and V_{oc} is open-circuit voltage. Finally, the power conversion efficiency of the solar cell is the proportion of the maximum power output to the input power (P_{in}) according to:

$$\eta = \frac{P_m}{P_{in}} = \frac{(FF \cdot J_{sc} \cdot V_{oc})}{P_{in}} \tag{10}$$

TABLE 1. The electrical parameters of the simulated PSC

Parameter	TiO_2	CH3NH3PbI3	Spiro-OMeTAD
Thickness (nm)	100	300	350
Eg (eV)	3.2	1.55	3.0
$\chi(eV)$	4.0	3.92	3.67
N_C (cm^{-3})	1×10^{19}	1×10^{19}	2×10^{19}
N_V (cm^{-3})	1×10^{19}	1×10^{17}	1×10^{19}
N_D (cm^{-3})	2×10^{18}	-	-
N_A (cm^{-3})	-	3×10^{15}	1×10^{18}
τ_n/τ_p (ns)	5/2	8/8	2/5

3. RESULTS AND DISCUSSION

Figure 1 depicts three-dimensional schematic views of perovskite solar cells with three different geometries used in our simulations. As illustrated in Figure 1(a), the simulated planar n-i-p PSC has a width of 250 nm and a

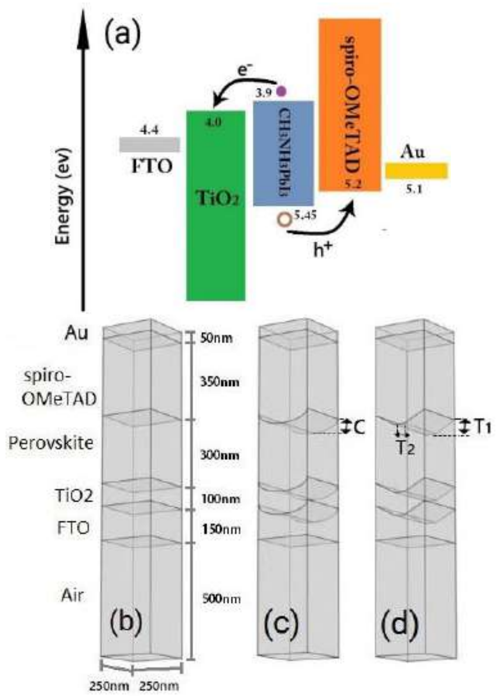


Figure 1. (a) Energy band diagram of the device. Schematic of the simulated PSC with (b) planar (reference cell) (c) concave and (d) trapezoidal structures. All simulated cells have the same dimensions

depth of 250 nm. The structure consists of Air (500 nm thick), FTO substrate (150 nm), TiO_2 as an ETL (100 nm), $CH_3NH_3PbI_3$ as an absorber layer (300 nm), Spiro-OMeTAD as a HTL (350 nm), and Au contact (50 nm). Figures 1(b) and (c) show the cells with the same dimensions on the nano-textured FTO substrates with the concave and trapezoidal groove shapes, respectively. In the concave architecture (Figure 1(b)), the concavity height is denoted by C and in the trapezoidal structure (Figure 1(c)) the height and the short base are denoted by T_1 and T_2, respectively. The energy band diagram of the device (Figure 1(d)) illustrates how photo-generated electrons and holes in the proposed designs are transported to contacts by TiO_2 and Spiro-OMeTAD as ETL and HTL, respectively.

First, to validate the correctness of our constructed model and selection of parameter values in our simulations (Table 1), we try to model and reproduce the experimentally obtained J-V characteristics of the experimental work from Barrit et al. [23] for planar structure (Figure 2).

After achieving a good agreement between the experimental and simulation results of the planar structure, we investigated the effects of using the proposed concave and trapezoidal-shaped FTO substrates on the performance of the PSCs. In order to achieve the optimum value for nanostructures, the height of the concave (C) and trapezoidal structures (T_1) are altered ranging from 40 nm to 70 nm. Meanwhile, the volume of the absorber layer and the thickness of the whole structure are kept unchanged. To clarify the effect of nanostructures shape and size on the absorption of the active layer we define the normalized absorption spectrum as follows:

$$\text{Normalized Absorption }(\lambda) = \frac{\text{Absorption of nanotextured substrate }(\lambda)}{\text{Absorption of planar substrate }(\lambda)}$$

The light absorption of planar structure (with and without Au) over the wavelength range of 300-800 nm is displayed in Figure 3(a), and the normalized absorption spectra in the absorber layer of the concave and

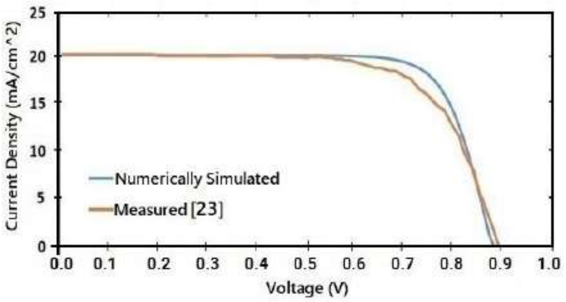


Figure 2. Comparison between the J-V characteristics of experimentally measured and numerically simulated planar perovskite solar cells

trapezoidal structures to their planar counterpart with different sizes are depicted in Figures 3(b-d). As can be seen in Figure 3(b), the maximum absorption improvement reaches around 32% at the wavelength of 560 nm. Although the best enhancement belongs to 70 nm concavity (C=70), the changes of absorption spectra with the height of concavity in most wavelengths are relatively small. Figure 3(c) shows the normalized absorption spectra of trapezoidal architecture with height (T_1) changing from 40 to 70 nm and constant short base of 50 nm (T_2=50 nm), results in a maximum absorption enhancement of 26% at a wavelength of 550 nm. Again, the changes of absorption spectra with trapezoidal groove height is negligible and it can be said that the absorption spectrum is almost independent of the nanostructures height. Figure 3(d) illustrates the normalized absorption spectra of trapezoidal architecture with a short base (T_2) starting from 40 nm to 100 nm, while the height is fixed at 50 nm (T_1=50 nm). It shows that the maximum absorption enhancement of 24% occurs at the wavelength of 380 nm for T_2=60 nm. In the wavelength range of 400-500 nm, we observe absorption increment for the case of trapezoidal structure with T_2=100 nm. To sum up, by integration of absorption values over the wavelength range of 300-800 nm, it can be said that upon using concave and trapezoidal structures with appropriate size the total absorption can be significantly enhanced compared to the planar structure.

Figure 4 illustrates the calculated carrier photo generation rate three-dimensional profiles (Equation (2)) at three different wavelengths of 300, 550 and 800 nm for planar, concave and trapezoidal solar cells. At λ=300 nm, most incident photons with high energy generate electron-hole pairs in TiO_2 and only a few photons can reach the front parts of perovskite and cause weak carrier photo-generation. Trapping of the light between nanostructures in active layer at this wavelength doesn't happen seriously since the most part of the light is absorbed at the ETL and couldn't reach the perovskite layer. At λ=550 nm (Figure 4(b)), however, the photons with lower energy cannot generate any carriers in TiO_2 due to its large bandgap.

Thus, we observe photo-generation of carriers only in $CH_3NH_3PbI_3$ layer especially in its front part near the ETL interface. The significant enhancement of the generation rate in the active layer of concave and trapezoidal structure in comparison with planar structure can be attributed to the light scattering mechanism occurs between the nanostructures wall. The result is in consistent with the normalized absorption spectra peaks (Figures 3(b) and (c)) around 550 nm. Finally, at λ=800 nm (Figure 4(c)) low-energy photons (1.5 eV) can penetrate to all layers of the cell, reach the gold contact and reflect back to the structure. This is related to the lower absorption coefficient of the absorber layer at such

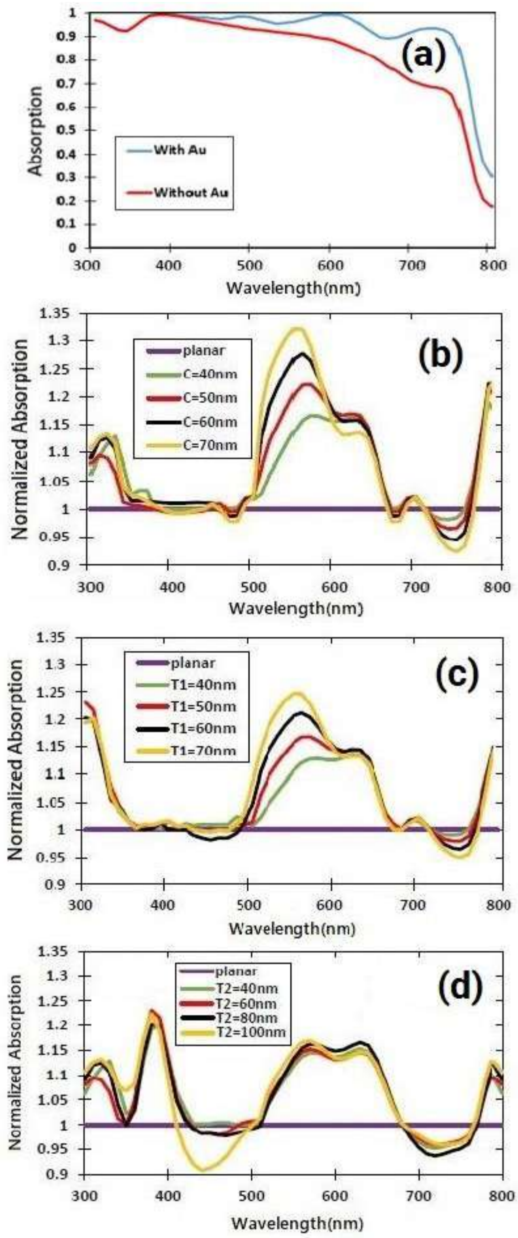


Figure 3. (a) Absorption spectrum of planar structure with and without Au contact layer. Normalized light absorption spectrum in the active area of the (b) concave (c) trapezoidal solar cells with changing the height of the nanostructures. (d) The same spectra for the trapezoidal solar cell with changing the short base size.

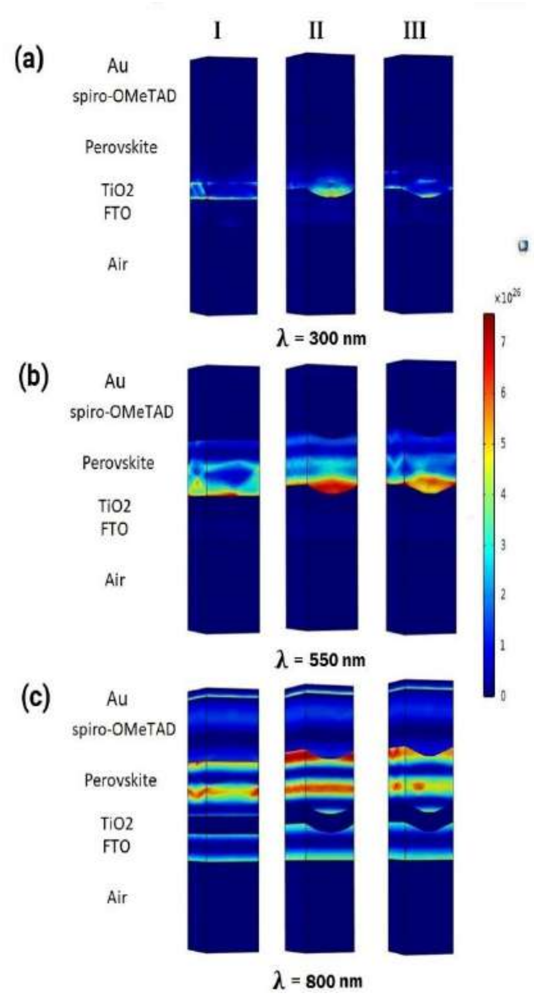


Figure 4. The calculated carrier photo-generation rate profiles ($G_{opt}(\lambda)$) for planar(I), concave (II) and trapezoidal (III) structures at three different incident wavelengths (a) 300 nm (b) 550 nm and (c) 800 nm.

high wavelengths. As can be seen in this wavelength, maximum G_{opt} occurs at the middle parts of $CH_3NH_3PbI_3$ layer and also near the interface of HTL due to the reflection and scattering of the light in the structure.

As explained before, by integration of $G_{opt}(\lambda)$ over the wavelength range of 300-800 nm, the total generation rate (G_{tot}) can be calculated (Equation (3)). Then, the obtained total charge density is used as the input of the semiconductor module to probe into electrical properties of the PSCs. Figure 5 displays the total photo-generated carrier profiles (G_{tot}) of the three proposed architectures under AM1.5 G solar spectrum illumination. Maximum total photo-generated carrier (G_{tot}) in perovskite happens in front parts of the layer near the interface of ETL.

Coupling the optical results with the electrical module, we can achieve J-V curve of the PSC structures as the most important electrical characteristics of the solar cell. Figures 6(a) and (b) depict IPCE spectrum and J-V curve of the solar cells with planar, concave and trapezoidal architecture, respectively. The concavity height (C), the trapezius height (T_1) and short base (T_2) in this simulation, all are set at 50 nm. Looking at Figure 6(b), it is conspicuous that the best J-V curve belongs to the concave structure. Such result was expected due to the higher absorption of concave cell relative to two other structures (Figure 3). Short circuit current density (J_{sc}), Open-circuit voltage (V_{oc}), Fill Factor (FF) and Power

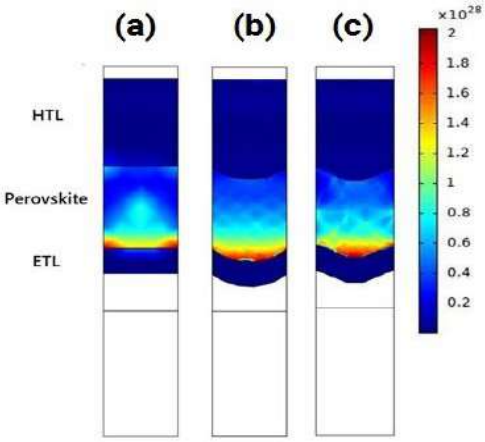


Figure 5. Total carrier photo-generation rate profiles (G_{tot}) in ETL, perovskite and HTL for (a) planar (b) concave and (c) trapezoidal structures over the wavelength range of 300-800 nm

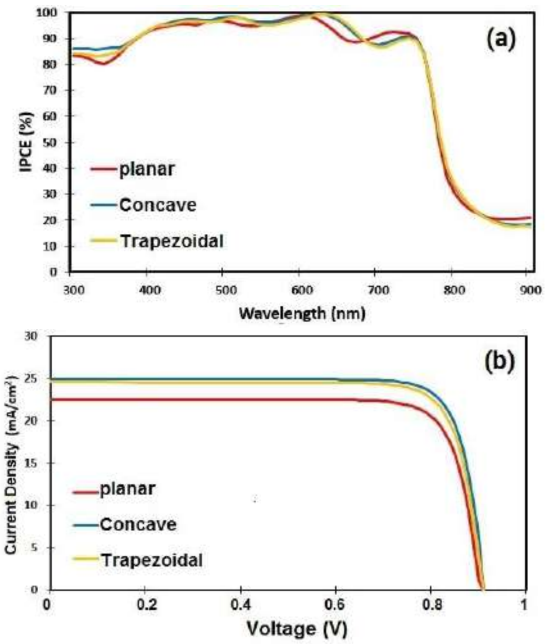


Figure 6. (a) IPCE and (b) Current Density-Voltage characteristics of the three proposed PSCs

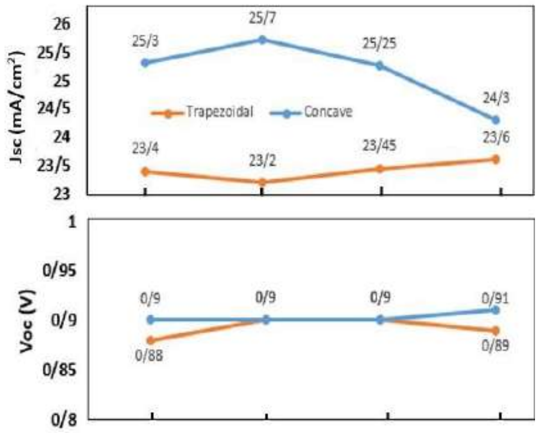
Conversion Efficiency (η) values as important electrical results obtained from the J-V curve are listed in Table 2. As can be observed in Figure 6 and Table 2, the open-circuit voltage (V_{oc}) is almost the same for all types of simulated PSC devices. This can be attributed to the same materials, thicknesses and contacts that we used in all structures in our simulations. Despite that, short circuit current density (J_{sc}) for devices with nano-textured substrate witnesses a substantial increase compared to their planar counterpart. The utmost improvement in short circuit current density belongs to the concave

TABLE 2. The electrical performance results of the planar, concave and trapezoidal PSCs in this work and nanotubular and corrugated structures in previous works

Structure	J_{sc} (mA/cm^2)	V_{oc} (V)	FF (%)	PCE (%)
Planar	20.06	0.88	75	13.16
The proposed concave	25.7	0.9	80	18.5
The proposed trapezoidal	23.2	0.9	79.5	16.6
Nanotubular [17]	24.29	1	62.7	15.2
Corrugated contact [18]	23.4	0.93	77.1	17.5

structure with J_{sc}= 25.7 mA/cm^2 which shows 28 and 10% increase in comparison with the planar and trapezoidal structures, respectively. The fill factor of the planar structure rises from 75 to 80% for concave and 79.5% for trapezoidal structures. This increase can be related to the enhanced interface area which leads to smaller series resistances between perovskite and carrier transfer layers. Regarding power conversion efficiency, the highest efficiency still belongs to the concave PSCs with 18.5% and trapezoidal PSCs with 16.6% which shows a significant surge compared to the planar PSC with around 13.2% efficiency.

Lastly, we investigate the effects of the nanostructures height on the electrical performance of PSCs. Figure 7 shows the variation of the J_{sc}, V_{oc}, FF and PCE as a function of concavity and trapezius height (C and T_1). The highest J_{sc} happens at C=50 nm for concave PSC and T_1=70 nm for trapezoidal structure. V_{oc} does not change significantly with the nanostructures height. It was predictable since difference between quasi-fermi levels under the solar illumination mainly depends on the materials of PSC, and the impact of nanostructures height is negligible. For both structures, the highest FF belongs to the 50 nm nanostructures height. Regarding the PCE, the best efficiency occurs at 50 nm height for concave structure (C=50 nm), reaching 18.5% efficiency. In



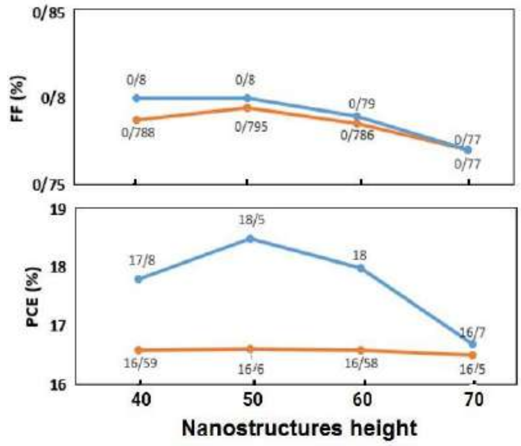


Figure 7. Variations of perovskite solar cell performance as a function of nanostructures height.

contrast to the concave structure, the PCE of trapezoidal PSC remains almost unchanged with the nanostructure height variations.

4. CONCLUSION

In this paper, the effect of using nano-textured substrates as light trapping architectures on improving the perovskite solar cells performance was investigated. Two different structures (concave and trapezoidal) of perovskite-based solar cells were designed, and different dimensions of the proposed structures were simulated to adopt the optimum values and enhance the overall efficiency. The analysis of the absorption in the active layer of the proposed structures normalized to the planar one over the wavelengths range of 300 to 800 nm demonstrated that the absorption enhances up to 32% in concave and 26% in trapezoidal structures at the wavelength of about 550 nm and at the height of 70 nm. In addition, we found that using concave and trapezoidal substrates can enhance the short-circuit current by around 28 and 15%, respectively. Accordingly, the power conversion efficiency was increased from 13.16% for planar PSC to 16.6% for trapezoidal, and 18.5% for concave structures at the height of 50 nm. It was shown that these improvements were related to the enhanced absorption, total carrier generation, and less series resistance in the proposed structures. It is thought that these achievements may open a new route for the fabrication of high efficiency perovskite solar cells through nano-textured substrates.

5. ACKNOWLEDGMENT

The authors acknowledge the funding support of Babol Noshirvani University of Technology, Iran through grant program No. BNUT/370557/99.

6. REFERENCES

1. Ke, W., Stoumpos, C.C., Zhu, M., Mao, L., Spanopoulos, I., Liu, J., Kontsevoi, O.Y., Chen, M., Sarma, D., Zhang, Y. and Wasielewski, M.R., "Enhanced photovoltaic performance and stability with a new type of hollow 3D perovskite {en} FASnI3." ***Science Advances,*** Vol. 3, No. 8, (2017), e1701293. doi: 10.1126/sciadv.1701293
2. Calvo, M. E., "Materials chemistry approaches to the control of the optical features of perovskite solar cells." ***Journal of Materials Chemistry A,*** Vol. *5*, No. 39, (2017), 20561-20578. doi: 10.1039/C7TA05666D
3. Fagiolari, L., and Federico, B., "Carbon-based materials for stable, cheaper and large-scale processable perovskite solar cells." ***Energy & Environmental Science,*** Vol. 12, No. 12, (2019), 3437-3472. doi: 10.1039/C9EE02115A.
4. Makableh, Y.F., Awad, I.A., Hassan, W. and Aljaiuossi, G., "Enhancement of the thermal properties of heterojunction perovskite solar cells by nanostructured contacts design." ***Solar Energy,*** Vol. 202, (2020), 204-209. doi: 10.1016/j.solener.2020.04.002
5. Wang, W., Winkler, M.T., Gunawan, O., Gokmen, T., Todorov, T.K., Zhu, Y. and Mitzi, D.B., "Device characteristics of CZTSSe thin-film solar cells with 12.6% efficiency." ***Advanced Energy Materials,*** Vol. 4, No. 7, (2014), 1301465. doi: 10.1002/aenm.201301465.
6. Khalate, S.A., Kate, R.S. and Deokate, R.J., "A review on energy economics and the recent research and development in energy and the Cu2ZnSnS4 (CZTS) solar cells: A focus towards efficiency." ***Solar Energy,*** Vol. 169, (2018), 616-633. doi: 10.1016/j.solener.2018.05.036.
7. Kojima, A., Teshima, K., Shirai, Y. and Miyasaka, T., "Organometal halide perovskites as visible-light sensitizers for photovoltaic cells." ***Journal of the American Chemical Society,*** Vol. 131, No. 17, (2009), 6050-6051. doi: 10.1021/ja809598r.
8. Zhang, F., Wang, Z., Zhu, H., Pellet, N., Luo, J., Yi, C., Liu, X., Liu, H., Wang, S., Li, X. and Xiao, Y., "Over 20% PCE perovskite solar cells with superior stability achieved by novel and low-cost hole-transporting materials." ***Nano Energy,*** Vol. 41, (2017), 469-475. doi: 10.1016/j.nanoen.2017.09.035.
9. Cahen, D., Zuo, C., Bolink, H.J., Han, H., Huang, J. and Ding, L., "Advances in perovskite solar cells." ***Advanced Science,*** Vol. 3, No. 7, (2016), 1500324. doi: 10.1002/advs.201500324.
10. Schulze, P.S., Bett, A.J., Winkler, K., Hinsch, A., Lee, S., Mastroianni, S., Mundt, L.E., Mundus, M., Würfel, U., Glunz, S.W. and Hermle, M., "Novel low-temperature process for perovskite solar cells with a mesoporous TiO_2 scaffold." ***ACS Applied Materials & Interfaces,*** Vol. 9, No. 36, (2017), 30567-30574. doi: 10.1021/acsami.7b05718.
11. Li, J. F., H. Y. Hao, J. B. Hao, L. Shi, J. J. Dong, H. Liu, and J. Xing. "Light trapping effect of textured FTO in perovskite solar cells." In IOP Conference Series: Materials Science and Engineering, Vol. 479, No. 1, 012046. IOP Publishing, (2019).
12. Saidaminov, M.I., Williams, K., Wei, M., Johnston, A., Quintero-Bermudez, R., Vafaie, M., Pina, J.M., Proppe, A.H., Hou, Y., Walters, G. and Kelley, S.O., "Multi-cation perovskites prevent carrier reflection from grain surfaces." ***Nature Materials,*** Vol. 19, No. 4, (2020), 412-418. doi: 10.1038/s41563-019-0602-2.
13. Tang, Z., Tress, W. and Inganäs, O., "Light trapping in thin film organic solar cells." ***Materials Today,*** Vol. 17, No. 8, (2014), 389-396. doi: 10.1016/j.mattod.2014.05.008.
14. Irandoost, R. and Soleimani-Amiri, S., "Design and analysis of high efficiency perovskite solar cell with ZnO nanorods and plasmonic nanoparticles." ***Optik,*** Vol. 202, (2020), 163598. doi: 10.1016/j.ijleo.2019.163598.

15. Abdelraouf, O.A., Abdelrahaman, M.I. and Allam, N.K., "Plasmonic scattering nanostructures for efficient light trapping in flat czts solar cells." In *Metamaterials XI***,** Vol. 10227, p. 1022712**. *International Society for Optics and Photonics***, (2017), doi: 10.1117/12.2265249.
16. Adinolfi, V., Peng, W., Walters, G., Bakr, O.M. and Sargent, E.H., "The electrical and optical properties of organometal halide perovskites relevant to optoelectronic performance." ***Advanced Materials,*** Vol. 30, No. 1, (2018), 1700764. doi: 10.1002/adma.201700764.
17. Abdelraouf, O.A. and Allam, N.K., "Towards nanostructured perovskite solar cells with enhanced efficiency: Coupled optical and electrical modeling." ***Solar Energy,*** Vol. 137, (2016), 364-370. doi: 10.1016/j.solener.2016.08.039.
18. Zandi, S. and Razaghi, M., "Finite element simulation of perovskite solar cell: A study on efficiency improvement based on structural and material modification." ***Solar Energy,*** Vol. 179, (2019), 298-306. doi: 10.1016/j.solener.2018.12.032.
19. Wenger, S., Schmid, M., Rothenberger, G., Gentsch, A., Gratzel, M. and Schumacher, J.O., "Coupled optical and electronic modeling of dye-sensitized solar cells for steady-state parameter extraction." ***The Journal of Physical Chemistry*** *C***,** Vol. 115, No. 20, (2011), 10218-10229. doi: 10.1021/jp111565q.
20. Schöche, S., Hong, N., Khorasaninejad, M., Ambrosio, A., Orabona, E., Maddalena, P. and Capasso, F., "Optical properties of graphene oxide and reduced graphene oxide determined by spectroscopic ellipsometry." ***Applied Surface Science,*** Vol. 421, (2017), 778-782. doi: 10.1016/j.apsusc.2017.01.035.
21. Ball, J.M., Stranks, S.D., Hörantner, M.T., Hüttner, S., Zhang, W., Crossland, E.J., Ramirez, I., Riede, M., Johnston, M.B., Friend, R.H. and Snaith, H.J., "Optical properties and limiting photocurrent of thin-film perovskite solar cells." ***Energy & Environmental Science,*** Vol. 8, No. 2, (2015), 602-609. doi: 10.1039/C4EE03224A.
22. Deceglie, M.G., Ferry, V.E., Alivisatos, A.P. and Atwater, H.A., "Design of nanostructured solar cells using coupled optical and electrical modeling**."** ***Nano Letters,*** Vol. 12, No. 6, (2012), 2894-2900. doi: 10.1021/nl300483y.
23. Barrit, D., Cheng, P., Tang, M.C., Wang, K., Dang, H., Smilgies, D.M., Liu, S., Anthopoulos, T.D., Zhao, K. and Amassian, A., "Impact of the Solvation State of Lead Iodide on Its Two-Step Conversion to MAPbI3: An In Situ Investigation." ***Advanced Functional Materials,*** Vol. 29, No. 47, (2019), 1807544. doi: 10.1002/adfm.201807544.

Persian Abstract

چکیده

در سال های اخیر ، استفاده از سلول های خورشیدی ترکیبی آلی غیر معدنی پروسکایت تحت طراحی پیشرفته و به دام اندازی نور، توجه زیادی را به خود جلب کرده است. در این مقاله، برای بررسی ویژگی های نوری و الکتریکی در سلول های خورشیدی پروسکایت با نانو ساختارها، از روش المان محدود سه بعدی استفاده شده است. با استفاده از لایه های اکسید قلع دوپینگ شده با فلور نانو بافت (FTO) ، ما دو معماری مختلف را پیشنهاد می‌کنیم که می‌توانند نور را در ابعاد نانومتر به دام بیندازند. دو ساختار پیشنهادی یعنی ساختارهای مقعر و ذوزنقه ای به منظور بررسی تأثیرات استفاده از لایه های نانوساختار بر عملکرد سلول های خورشیدی پروسکایت با ساختار مسطح مقایسه می‌شوند. تجزیه و تحلیل نوری نشان می‌دهد که استفاده از ساختارهای مقعر و ذوزنقه‌ای بهینه می‌تواند جذب نور را در طول موج 550 نانومتر به ترتیب تا 32 و 26 درصد افزایش دهد. شبیه سازی های الکتریکی نشان داده اند که علاوه بر افزایش تولید حامل های کل ، حامل های تولید شده را می توان به طور موثر در نانو ساختار پیشنهادی بدست آورد. بر این اساس ، جریان اتصال کوتاه از 20 میلی آمپر برای ساختار مسطح به 25.7 میلی آمپر برای مقعر و 23.2 میلی آمپر برای سلول خورشیدی پروسکایت با معماری ذوزنقه ای افزایش یافته است. پس از تجزیه و تحلیل ارتفاعات مختلف و با اتخاذ مقادیر بهینه ، بازده تبدیل نیرو برای سلول های خورشیدی پروسکایت با ساختارهای مقعر و ذوزنقه نسبت به ساختار مسطح 5.5٪ و 3.5٪ افزایش قابل ملاحظه‌ای داشته است. این پیشرفت های قابل توجه که توسط مدل نوری و الکتریکی همراه نانو ساختار مورد تجزیه و تحلیل قرار گرفته است ، می‌تواند راه را برای مطالعات بیشتر برای ساخت سلول های خورشیدی پروسکایت با بازدهی بالا با لایه های نانو ساختار به عنوان روش به دام اندازی نور فراهم کند.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 881-890

International Journal of Engineering

Journal Homepage: www.ije.ir

Voltage Stability Improvement in Optimal Placement of Voltage Regulators and Capacitor Banks Based on FSM and MMOPSO Approach

M. Shafiee[a], M. Amirahmadi*[b], M. Farzinfar[c], A. Lapthorn[a]

[a] *Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand*
[b] *Department of Electrical and Electronic Engineering, Semnan Branch, Islamic Azad University, Semnan, Iran*
[c] *School of Engineering, Damghan University, Damghan, Iran*

PAPER INFO

Paper history:
Received 12 June 2020
Received in revised form 11 November 2020
Accepted 01 January 2021

Keywords:
Allocation
Capacitor Bank
Multi-objective Particle Swarm Optimization
Voltage Regulator
Voltage Stability Enhancement

ABSTRACT

Installation of Shunt Capacitor Banks (SCBs) and Voltage Regulators (VRs) within distribution system is one of the most effective solutions in reactive power control for improving the voltage profile and reducing power losses along the feeder. However, the presence of the VRs can deteriorate the Voltage Stability Margin (VSM) in distribution feeders. To address this issue, this paper proposes a multi-objective programming model for the simultaneous optimal allocation of VRs and SCBs in the distribution network to improve the voltage profile and to minimize power losses and installation costs. In the proposed model, a Voltage Stability Index (VSI) is considered to prevent voltage instability during SCBs/VRs allocation. A new Modified Multi-Objective Particle Swarm Optimization (MMOPSO) algorithm which includes a dynamic inertia weight and mutation operator is proposed to obtain the optimal solutions as a Pareto set. Thereinafter, a Fuzzy Satisfaction Method (FSM) determines the optimal solution. A practical long radial distribution feeder has been employed to demonstrate the efficiency and efficacy of the proposed model along with a comparison between the proposed MMOPSO and the original MOPSO.

***doi**: 10.5829/ije.2021.34.04a.14*

1. INTRODUCTION

Long feeders are prevalent in practical distribution grids as they deliver energy to scattered consumers and, mainly, areas with low population density. Considerable voltage sag and power losses are common issues of such feeders. Voltage drop is one of the major contributors to increases in network losses and feeders' operation at full load capacity is often impossible due to the voltage drop. In some cases, this issue leads to reducing feeder loading to less than 10% of feeder rated capacity [1]. In contrast, the design and construction of new HV substations and MV networks close to the demand side are not feasible for reasons such as low demand or due to economic limits of investment. Therefore, it is necessary to find alternative solutions to tackle these problems and ensure stability, reliability, and quality of the electric power supply [2].

To date, many different approaches have been proposed to resolve the issues mentioned above. Installing Shunt Capacitor Banks (SCB), as the reactive power compensators, is a well-known and common solution [3-5]. Further, voltage control can be effectively achieved using in-line automatic voltage regulators (VR), which consists of an autotransformer fitted with a tap-changing mechanism [6-8]. VR's advantages, such as controlling the voltage magnitude within standard ranges, have persuaded utilities to utilize it in distribution grids [9]. These methods each have their own advantages and disadvantages. For example, the main advantage of utilizing SCB is its simplicity and low implementation cost. However, in the case of overvoltage situations, which may occur in feeders with a high penetration level of Distributed Generation (DG) or light-load conditions, VRs are more capable of controlling conditions compared to SCB [10-12]. Although, inappropriate

*Corresponding Author Institutional Email: *m.amirahmadi@semnaniau.ac.ir* (M. Amirahmadi)

Please cite this article as: M. Shafiee, M. Amirahmadi, M. Farzinfar, A. Lapthorn, Voltage Stability Improvement in Optimal Placement of Voltage Regulators and Capacitor Banks Based on FSM and MMOPSO Approach, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 881-890

installation of VRs in long distribution feeders might have devastating effects on voltage stability, and it may lead to voltage collapse.

The benefits of voltage control and loss reduction can best be realised with the optimal allocation of this type of equipment. In fact, inappropriate placing of SCBs or VRs along with improper sizing, may lead to even higher system losses or voltage violations.

Due to the complexity of the problem of locating voltage control devices, like SCBs or VRs, the use of optimization techniques, with goals such as loss reduction and voltage improvement, is widespread in the literature [6, 11-16]. Articles published in this field can be easily divided into two general categories:

The first group discusses the allocation of various control devices, such as SCBs, VRs, DGs, and the goals that must be met in the allocation (such as reducing losses and emission and improving network voltage) [17]. Further, in the second category, different optimization algorithms are presented, knowing that the optimal allocation problem with different objectives is a complex problem. For example, extensive literature [4, 18-21] address SCBs allocation in the distribution network for voltage control and energy losses minimization. The optimal allocation of SCBs based on two optimization techniques, i.e., water cycle algorithm (WCA) and grey wolf optimizer (GWO) have been addressed [22]. Hocks et al. [8] have described in detail the performance of the VR in improving the power quality of the network. Optimal placement and sizing of VRs in a radial distribution network have been discussed in literature [13, 14]. The literature [7, 23, 24] have suggested the simultaneous allocation of SCBs/VRs in order to take advantage of both devices. Further, in some research, hybrid placement and sizing of SCBs/DGs to improve power losses and voltage profile of the distribution network and minimization of investment cost have been considered [11, 25-27]. To enhance the real Egyptian distribution system's performance, optimal placement of the combined SCBs, DGs, and VRs has been introduced in literature [28-30] based on the PSO algorithm. Although literature [28] provides a comprehensive overview of system performance improvement techniques, it has not examined the impact of placement on network stability.

Although the methods presented in the articles mentioned above for SCBs/VRs allocation lead to grid performance improvement, one of their drawbacks is ignoring the network voltage stability during the optimal allocation. The voltage collapse phenomenon is critical in long feeders, which has been the subject of many papers [1, 31-33]. It is a fact that VRs increase injected reactive power to the feeder, which results into lower lag power factor, and consequently, lower Voltage Stability Margin (VSM). In contrast, SCBs as reactive power sources, can improve the VSM. Due to the different effects of these two elements on the VSM, the voltage stability studies should be factored in the simultaneous installation of VRs and SCBs. Given the above shortcoming, this paper introduces a voltage stability index to prevent the feeder's voltage collapse while optimal allocating of the SCB/VRs.

Franco et al. [7] have carried out research by using the Mixed Integer Linear Programming (MILP) model for optimal placement of SCBs/VRs in a distribution system. Tolba et al. [19] hybridization of Particle Swarm Optimization besides a Gravitational Search Algorithm (PSOGSA) is suggested for solving the optimal allocation of SCBs. Oliveira et al. [34] have applied Mixed Integer Non-Linear Programming (MINLP) model to obtain the optimum size of the SCBs. Evolutionary and nature inspired techniques, like Genetic Algorithm [24, 35] or Particle swarm optimization (PSO) [23], Harmony Search Algorithm (HSA) [36], Water Cycle Algorithm [11], have also extensively used for solving allocation problems due to their specificity in solving the optimization problem.

As already discussed, the problem of SCBs/VRs is a complicated multi-objective problem, which due to being trapped in local optima and the possibility of premature convergence, the simple or original class of heuristic methods may not be sufficient to find the optimal solution. In this regard and as the second contribution, a novel Modified Multi-Objective PSO (MMOPSO) is proposed, which obtains optimal Pareto set as optimal solutions. Thereinafter, the Fuzzy Satisfaction Method (FSM) is employed to determine the best optimum solution. The main features and contributions of this research study are highlighted as follows:

- Optimal allocation of the SCBs/VRs to enhance technical and economic issues of distribution systems.
- Introducing a voltage stability index (VSI) as one of the objective functions during allocation, in addition to power losses, voltage deviation, and SCBs/VRs installation costs. The advantage of considering VSI is to prevent the voltage collapse of long feeders after the VRs installation.
- Proposing a Modified Multi-Objective PSO (MMOPSO) algorithm to find optimal solutions for optimal allocation of SCBs/VRs program.
- Applying a Fuzzy Satisfaction Method (FSM) for determining the optimum solution among the non-dominated Pareto set.
- Applying the proposed method to a real radial distribution system.

2. PROBLEM FORMULATION

This paper introduces a multi-objective programming model to optimally allocate the SCBs/VRs within a long feeder. According to the model presented, SCBs/VRs are positioned along the feeder to optimally improve all operating criteria such as losses and voltage indices. In this section, the objective functions (OFs) and related constraints are introduced.

2. 1. Power Losses Minimization of the total feeder losses after installation of SCBs/VRs is technically considered as an objective function in optimal allocation of SCBs/VRs [37, 38]

$$\min P_{Loss} = \sum_{ij=1}^{N_L} R_{ij} \times I_{ij}^2 \tag{1}$$

where Rij stands for the resistance of each line section within the feeder, and I_{ij} indicates current flow through them.

2. 2. Voltage Deviation In the second OF, minimization of total voltage deviation in 20kV feeders is considered as follows:

$$\min \sum_{n=1}^{N_b} |V_{ref} - V_n| \tag{2}$$

In this paper, the desired and reference voltage (Vref) has not been considered a constant value, e.g., 1 p.u.; in contrast, it has been assumed to be a range of standard voltages. In other words, the voltage magnitude at each terminal must be maintained within specified limits as follows:

$$0.95 \le V_{ref} \le 1.05 \tag{3}$$

As per the above constraint, voltage deviation happens whenever the voltage of any terminal (V_n) along the feeder is out of this allowable range. One of the benefits of using such a function is to avoid unnecessary installation of SCBs/VRs to reduce the total costs. In other words, by considering 1 p.u. for V_{ref}, the optimization method would result in higher capacity or number of SCBs/VRs as it tries to reduce deviated voltage from V_{ref}.

2. 3. Installation Cost of SCBs/VRs Due to the limited budget of distribution companies, allocating SCBs/VRs are not possible without considering the economic parameters. Thus, one of the essential objective functions that are generally considered in the optimization problem is an economic OF. This OF aims to optimally determine the number of allocated SCBs/VRs from the viewpoint of cost.

$$\min C = \sum_{i=1}^{N_{cap}} Cap\_Kvar_i \times Cost_{cap} + N_{VR} \times Cost_{VR} \tag{4}$$

As stated in Equation (4), the total cost imposed by the SCBs/VRs installation depends on their capacity/number. According to price inquiry from different suppliers, two values of 4.5 US$/kVAR and US$45000 have been considered, respectively, for evaluation of $Cost_{cap}$ and $Cost_{VR}$ terms in Equation (4). In this paper, the capacity of each SCB is considered 200KVAR. Due to the limited budget in implementing SCBs/VRs, the maximum number of SCBs is limited to 10 and 20 (In two different studied scenarios in simulation results), and the number of VRs is limited to 3.

2. 4. Voltage Stability Index Unlike during normal operating conditions where a slight increase in load causes a slight decrease in voltage, if the feeder loading exceeds a certain value, then an increase in load will result in a rapid decrease in voltage and, eventually, voltage collapse. Installing the SCBs/VRs without consideration of this issue, despite its other advantages, can bring the network closer to the voltage instability point. Conversely, properly installation of SCBs/VRs can save the network from voltage collapse by changing the flow of active/reactive power.

Given the above description, and considering the ever-growing load demand in distribution networks, it is imperative to consider the stability issue along the feeder while allocating the SCBs/VRs. This issue, to the best of the authors' knowledge, has been overlooked in the literature. Hence, a voltage stability index (VSI), which has been derived based on the research reported in literature [32], is presented that should be minimized in the optimal allocation problem.

$$\min \ VSI = 4\left[\left(x_{eq}P_{Deq} - r_{eq}Q_{Deq}\right)^2 + x_{eq}Q_{Deq} + r_{eq}P_{Deq}\right] \tag{5}$$

Since the practical distribution network consists of many lines and laterals, the stability index is obtained based on the single-line method for reducing a distribution network and extraction of the equivalent Thevenin's parameters.

In this equation, PDeq and QDeq are total real and reactive loads in the distribution network. Also, equivalent resistance (req) and equivalent reactance (xeq) for a single line are defined as follows:

$$r_{eq} = \frac{\sum P_{Loss}}{\left\{\left(P_{Deq}+\sum P_{loss}\right)^2+\left(Q_{Deq}+\sum Q_{loss}\right)^2\right\}} \tag{6}$$

$$x_{eq} = \frac{\sum Q_{Loss}}{\left\{\left(P_{Deq}+\sum P_{loss}\right)^2+\left(Q_{Deq}+\sum Q_{loss}\right)^2\right\}} \tag{7}$$

where

$$P_{Loss} = R_{ij}\frac{{P_{ij}}^2+{Q_{ij}}^2}{{V_n}^2} \tag{8}$$

$$Q_{Loss} = X_{ij}\frac{{P_{ij}}^2+{Q_{ij}}^2}{{V_n}^2} \tag{9}$$

where (P_{ij}, Q_{ij}) are the active/reactive power of each line, which are altered by utilizing SCBs/VRs along the feeder.

It should be noted that theoretically, in a stable system, the VSI index has to be below 1. If the feeder is loaded beyond the critical limit, at this circumstance, the voltage will collapse.

In addition to the mentioned objective functions, distribution system constraints such as a line's loading capacity are considered in the problem.

3. SOLUTION APPROACH

The problem of SCBs/VRs allocation is a complicated multi-objective problem that should be solved via a multi-objective optimization algorithm. Therefore, heuristic methods, like PSO and GA, have been commonly used in the literature [8, 13, 36, 38]. However, the original or simple class of these kinds of methods are susceptible of being trapped into local optimum and also premature convergence and may not be sufficient to find the optimum solution, especially within practical distribution networks that have numerous long feeders. In this regard, a novel modified Multi-Objective PSO for finding an optimal Pareto set of non-dominated solutions is proposed as the paper's second contribution. Moreover, the FSM technique for determining the optimum solution among the Pareto set is presented.

3. 1. Modified Multi-objective PSO (MMOPSO)

Multi-objective problem optimization is a class of complex problems with objective functions, which can be incomparable or contradictory. In these problems, it is impossible to find a single global optimum solution. Contrary to the single-objective optimization scenario, there is an optimal set of solutions or alternatives, called the Pareto optimal set. Expert analysis and trade-off can, therefore, describe the optimal solution among the Pareto set. The aim of the multi-objective model proposed in the present research is to find the optimal size and site of SCBs and the location of VRs along the distribution feeders.

However, solving the MOPs with mathematical or linear programming methods is difficult due to the necessity of optimizing several objective functions. Metaheuristic algorithms, such as PSO, are therefore suitable because of their ability to synchronously search for multiple Pareto optimal solutions and perform better global exploration and local exploitation of the search space. PSO is a population-based optimization strategy inspired by bird flocking or fish schooling social behaviour [39] and extended by authors of [40, 41] as multi objective PSO. Consider X_i $(X_{i,1}, X_{i,2}, \dots, X_{i,d})$ denotes an n-dimensional decision variable vector which moves with a velocity V_i $(V_{i,1}, V_{i,2}, \dots, V_{i,d})$. The particles positions are restored as nondominated vectors in the repository (REP). The historical record of a particle's best solutions is used for storing non-dominated solutions. Each particle is associated with its best solution achieved, X_{best-i} $(X_{best-i,1}, X_{best-i,2}, \dots X_{best-i,d})$, which is defined by its own best performance in the swarm. Further, each particle in its movement selects a member of REP randomly as its leader (X_i^{REP}). In this paper, the selection of the leader from the RES is done based on the hypercube method and applying the Boltzmann and roulette-wheel selection algorithm [41].

The particle's new position is controlled by updating its position attributes and velocity (10 and 11). A modified version of the classical PSO (MPSO) based on literature [42] has been used in this paper to ensure the optimal solution in the allocation problem. The third term of the Equation (10) is a randomly applied intermediate crossover in the MPSO to prevent particles from becoming lazy in the swarm after a while.

$$V_i^{iter+1} = W \times V_i^{iter} + c_1 \times rand \times \left(X_{best-i}^{iter} - X_i^{iter}\right) + c_2 \times rand \times \left(X_i^{REP} - X_i^{iter}\right) + c_{1c} \times rand \times \left(X_i^{iter} - X_{i_{not-best}}^{iter}\right) \tag{10}$$

$$X_i^{iter+1} = X_i^{iter} + V_i^{iter+1} \tag{11}$$

where W is the inertia weight, c_1 and c_2 are cognitive and social acceleration coefficients, and $X_{i_{not-best}}^{iter}$ reflects the worst particle experience.

Two modifications are implemented in this paper to solve the proposed multi-objective model and to improve the MPSO 's efficiency in finding the global optimum. These modifications improve the algorithm's convergence capability and searchability. The following describes these modifications, respectively.

3. 1. 1. Dynamic Inertia Weight The particles inertial behaviour causes a partial restriction of the particle velocity variations so that the particles from the search space don't change their direction quickly to the best swarm experience; thus, fast convergence of the algorithm is prevented. At the start of random search algorithms, such as PSO, an exploration or global search is required to find the optimal search space, so the diversity of the population in the initial iterations should be preserved. The particles should also explore the entire search space that is met by selecting a relatively high value for inertia weight. Setting too high values for inertial weight causes a problem in that the algorithm in the final iterations cannot correctly converge to the X_{Gbest} experience. Thus, the inertia weight should be selected as a balance between exploration and exploitation so that the algorithm addresses both issues. In this paper, a dynamic inertia weight factor is introduced as follows to resolve these issues and to maintain the balance between exploration and exploitation:

$$W = \left(\frac{iter_{max} - iter}{iter_{max}}\right) \tag{12}$$

As per (12), with increasing repetitions, gradually and dynamically, the weight of inertia decreases as the algorithm converges to the optimal point and the best group experience. Besides, the inertia weight has characteristic randomness to maximize the particle variety.

3. 1. 2. The Proposed Mutation Mechanism In a problem with a vast search space, like the problem of optimal SCBs/VRs allocation, in which every node along the feeder could be a candidate, it is possible that the initial population is far from the optimum solution. Under such a case, the MPSO 's exploration capability degrades, and therefore the particles move rapidly to a false Pareto front, which may be a local optimum under global optimization, leading to premature convergence. A mutation operator (Equation (13)) is introduced to tackle this issue, which causes particles to move in different directions during the optimization process and enhance the algorithm's exploration.

$$X_i^{iter} = X_i^{iter} + W \times \frac{1}{\pi\gamma}\left[\frac{\gamma^2}{\left(X_i^{iter}-\theta\right)^2+\gamma^2}\right] \tag{13}$$

The proposed operator is multiplied by the inertia weight (W), which in the beginning increases the step size of the mutation and thus increases the opportunity to search for new areas. In this way, the explorative behaviour of the algorithm is improved. By comparison, as the current best solution reaches an optimum solution in subsequent iterations, the mutation's step size is reduced to improve the accuracy of convergence. The mutation, on the other hand, should be done at random; thus, when a particle is selected for mutation, a random disturbance is applied to its current position. The probability in this paper is determined on the basis of the Cauchy distribution function, which is multiplied by W. The scale parameter (γ) and location parameter (θ) are set to 0.5 and 0, respectively, in the distribution function.

3. 2. Fuzzy Satisfaction Method (FSM) The equations called membership functions are used to define the fuzzy sets. These functions show membership level in certain fuzzy sets using values from 0 to 1 [43]. The membership value '1' presents compatibility totally, while the number '0' means full incompatibility with the set.

In this study, fuzzy sets are defined for all the objective functions discussed in section 2. We assume that all objective functions' maximum and minimum values can be defined based on the base case scenario and other aforementioned technical constraints. By taking account of the individual minimum and maximum values of each objective function, the membership function $\mu(R_i)$ for each objective function in section 2 can be determined in a subjectively manner as expersion (14), Where R_{min}^i and R_{max}^i are the minimum and maximum values of ith objective function in which the solution is expected.

$$\mu(R_i) = \begin{cases} 1 & R_i < R_i^{min} \\ \dfrac{R_i^{max} - R_i}{R_i^{max} - R_i^{min}} & R_i^{min} \le R_i \le R_i^{max} \\ 0 & R_i > R_i^{max} \end{cases} \tag{14}$$

For calculating the membership functions, the minimum and maximum values of objective functions must be defined first. The maximum values of losses and total voltage deviation are set according to the data resulting from the base case simulation at the test case feeder. This is because the values for losses and total voltage deviation in the base case are the worst between Pareto solutions, and the optimization methods are aimed to reduce these values. The minimum values for these two objective functions theoretically could reach the value '0'. In a stable system, the VSI index has to be below 1, thus the maximum and minimum values of this objective function are defined '1' and '0', respectively. Due to the limited budget in implementing SCBs/VRs, the maximum numbers of these elements are limited to the mentioned values in section 2. Therefore, the maximum values of the installation cost of SCBs/VRs are calculated as the maximum number of SCBs/VRs multiplied by the stated prices in section 2. The minimum values of the installation cost for SCBs/VRs are zero.

The membership functions' value indicates how much (in scale from 0 to 1) a solution is satisfying the objective. The minimum value of all membership functions for a specific combination represents the optimality value of the combination. Therefore, a combination with a larger minimum value of membership functions is more favourable since it can lead to more objective functions tending to their individual optimum values. Thus among all possible optimum solutions, one should seek a combination for which the minimum value of all the membership functions is maximum. Hence for a multi-objective optimization problem with N objective functions, the following index (φ) can be calculated for every Pareto solution in the repository.

$$\varphi \equiv \mathrm{Max}\{\min(\mu(R_i))\} \qquad i = 1, \dots, N \tag{15}$$

According to the relation (15), the Pareto set's optimal solution would be the Pareto solution for which φ is the maximum.

4. SIMULATION RESULTS

To verify the performance of the proposed algorithm in obtaining the optimal placement and capacity of SCBs/VRs, a practical 20 kV distribution network located in Semnan (Figure 1) is used as a test system. For verification of the proposed approach's efficacy, a long feeder (about 180 km) is selected for the allocation of SCBs/VRs. Feeder information in the base case (where there are no SCBs/VRs within the feeder) are shown in Table 1. In Figure 1, all feeders are shown in different colours. The main advantage of choosing a practical test system includes:

A large and vast power system is more complex, so if the proposed algorithm could obtain suitable solutions, it can be assured that it also works for smaller systems.

It was possible to access hourly load information of the test system over the past year, which is one of the algorithm implementation requirements.

The proposed MMOPSO starts optimizing the problem with a population of 60 particles, a repository size of 15 particles; and finally, the algorithm stops after 100 iterations. It should be noted that, in the proposed MMOPSO, the related parameters are set based on the values obtained in the literature [44, 45]. To evaluate the effectiveness of the simultaneous allocation of SCBs/VRs, two scenarios are studied for the purpose of discussion.

In the first scenario, the optimal simultaneous allocation of SCBs/VRs based on the proposed formulation is obtained, while, in the second one, the optimum allocation of SCBs without considering the VRs is done to evaluate the effect of the VRs on the network operation. In both scenarios, the implementation of the proposed MMOPSO in finding optimum solutions is evaluated in contrast to the original multi-objective PSO.

TABLE 1. Power flow data of studied feeder in the base case

Active Power (MW)	Reactive Power (Mvar)	Nominal Voltage (KV)	Input Current (A)	Losses (MW)
3.53	1.85	20	115	0.796

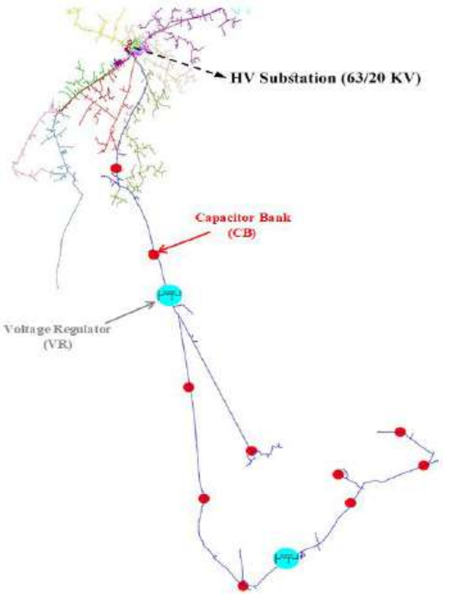


Figure 1. Location of SCBs/VRs in scenario 1 in Semnan distribution network in DIgSILENT

4. 1. Scenario 1. Simultaneous Allocation of SCBs/VRs As already mentioned, in this scenario, the optimal allocation of SCBs/VRs and their impacts on improving the operation of the grid is assessed. The repository consists of 5 Pareto optimal solutions for solving the problem, which are reported in Table 2. The optimal location of the allocated VRs and SCBs based on the obtained results (for the solution with the maximum φ in Pareto set - solution 4 in Table 2) are also depicted in Figure 1. For the sake of comparison, some operating indices and objective functions in the base case are also inserted in this table. In all Pareto solutions, the optimum number of VRs was 2. So, the number and capacity of SCBs, as the difference between Pareto solutions, have been inserted in this table instead of installation cost.

As per the results of Table 2, in the simultaneous allocation of SCBs/VRs, the maximum number of SCBs that can be placed is 10. Nevertheless, all five reported solutions improve network performance in terms of voltage improvement and loss reduction compared to the base case. However, none of the solutions is dominated. The FSM method is employed to determine the best solution in this table. By calculating index φ, it can be found that solution 4 satisfies Equation (15), and it is the best solution in the Pareto set. The value of φ in this solution is 0.2977.

The obtained results verify the performance and effectiveness of the proposed method, as follows:

- In all of the Pareto solutions, all of the objective functions' values are more optimal than the base case.
- The individual values of objective functions (except voltage deviation index) in solution 4 are not the best between all solutions, but calculating index φ shows that solution 4 is the best. In fact, there is a compromise between different objective functions in this solution.
- The losses amount in solution 4 is decreased to 0.559 MW, as given in Table 2. However, it is not the minimum value in all solutions, which indicates the non-domination of different solutions in the Pareto set. The best solution from this viewpoint is solution 3.
- Voltage deviation happens whenever the voltage of any terminal (V_n) along the feeder is out of the allowable range. As stated in Equation (2), the total voltage deviation equals the summation of the voltage deviations at all feeder busses. As illustrated in Table 2, the total voltage deviation of the feeder in scenario 4 considerably decreases from 253.53 p.u. to 0 p.u. In this regard, the minimum voltage value is increased to 0.954 p.u., which indicates that the voltage of all terminals is controlled within the permissible range, compared to the base case.

To compare the network operation conditions, Figure 2 and Figure 3 illustrate the voltage profile of the test case

TABLE 2. The optimal allocation of SCBs/VRs (scenario 1)

Optimum Solution	MMOPSO Results					Original MOPSO Results				
	Losses (MW)	Total Voltage Deviation (p.u)	VSI	Total installed SCBs (Mvar)	Number of SCBs	Losses (MW)	Total Voltage Deviation (p.u)	VSI	Total installed SCBs (Mvar)	Number of SCBs
Base case	0.796	253.53	0.632	0	0	0.796	253.53	0.632	0	0
1	0.549	0	0.563	1.67	10	0.541	6.98	0.554	1.95	10
2	0.533	11.84	0.547	1.51	10	0.554	24.58	0.576	1.55	9
3	0.531	19.00	0.589	1.46	10	0.569	12.25	0.561	2	10
4	0.559	0	0.556	1.39	10	0.576	42.35	0.581	1.2	8
5	0.589	66.70	0.567	1.1	8	0.552	5.23	0.572	1.8	10

feeder in the base case and scenario 4, respectively. It is clear that the feeder voltage profile is improved under the simultaneous allocation of SCBs/VRs conditions.

- Another salient feature of the proposed methodology is to increase the voltage stability criterion, following optimal SCBs/VRs allocation. The results show that the VSI index is strengthened relative to the base case, which allows for more loading of the feeder if needed. It should be noted that when the VSI approaches one, it indicates that the system is close to the voltage collapse condition. It is worth noting that although the power losses in solution 1 is lower than 4, but it is the VSI in scenario 4 that is better than 1, resulting solution 4 is the best in this scenario.
- In Table 2, the optimal solutions obtained by the original MOPSO are also reported, comparing them with the results of the proposed MMOPSO. The obtained results show that the value of φ in the original MOPSO is 0.2764, which is lower than the φ in MMOPSO case. It demonstrates the feasibility and efficiency of the MMOPSO algorithm in terms of allocation of SCBs/VRS, and so improved operating conditions.

4. 2. Scenario 2- Allocation of SCBs

To examine the effect of the VRs on improving feeder performance, in the second scenario, only the optimal allocation of SCBs is discussed. In other words, in this case, the VRs are eliminated and the MMOPSO and original MOPSO methods are applied to optimize the site and size of SCBs only. The maximum number of allocated capacitors is increased from 10 to 20 in this scenario to distinguish the results better.

The 5 Pareto optimal solutions are reported in Table 3. The obtained results can be summarised as follows:

- The MMOPSO allocates the SCBs along the feeder in order to satisfy the objective functions of the voltage control and loss reduction; however, it has not been able to adjust the feeder voltage compared to scenario 1 optimally.
- Depicted in Table 4, the φ values for discussed scenarios indicate that the φ values in scenario 1 for both algorithms are more optimal than the ones in scenario 2.
- Based on statistics seen in Table 4, solution 1 in MMOPSO is the best in this scenario, where the total capacity of 2.61 MVar capacitors are allocated, and the minimum voltage of feeder terminals is obtained 0.922 p.u., which is outside the permissible range. Also, in this case, the total voltage deviation is calculated by 140.54 p.u., which is not acceptable.

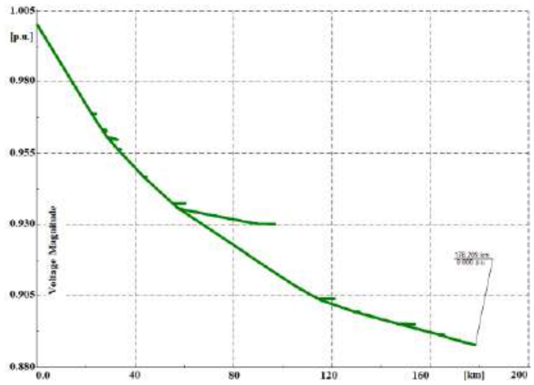


Figure 2. Voltage profile of test case feeder in the base case

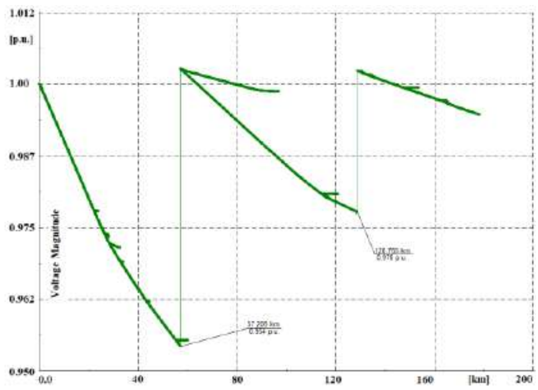


Figure 3. Voltage profile of test case feeder in the scenario 1

TABLE 3. The optimal allocation of SCBs (scenario 2)

Optimum Solution	MMOPSO Results					Original MOPSO Results				
	Losses (MW)	Total Voltage Deviation (p.u)	VSI	Total installed SCBs (Mvar)	Number of SCBs	Losses (MW)	Total Voltage Deviation (p.u)	VSI	Total installed SCBs (Mvar)	Number of SCBs
Base case	0.796	253.53	0.632	0	0	0.796	253.53	0.632	0	0
1	0.591	140.54	0.508	2.61	17	0.754	102.34	0.592	3.12	18
2	0.788	87.72	0.601	3.46	19	0.651	194.76	0.536	0.98	13
3	0.667	207.38	0.544	0.78	12	0.642	189.23	0.538	1.24	14
4	0.636	191.78	0.530	1.16	10	0.602	153.75	0.522	2.85	19
5	0.650	199.74	0.536	1.01	12	0.612	132.24	0.541	2.91	18

- Analyzing the average values of VSI in both scenarios reveals that implementing VRs deteriorates the voltage stability status in distribution grids, especially in long feeders. Therefore, voltage stability indices should be taken into account as one of the critical objectives in allocating VRs.

To conclude, it is clear that the allocation of capacitors alone cannot solve the problem of operating long feeders. On the other hand, it is not technically correct to install too many SCBs along the feeder as placing capacitors in a system could cause harmonic resonance and/or switching transients and may cause equipment damage due to high voltage, excessive thermal problems, and current circulation between the capacitor and the system. Finally, the results of Table 3 confirm that the proposed MMOPSO algorithm is more successful and more reliable in finding the optimal solution than the original MOPSO.

The main finding from the obtained results can be summarized as follows:
In long feeders without use of VRs and only with the use of SCBs, it is not possible to adjust the voltage of the feeder properly. In fact, in the medium voltage networks, due to the predominance of R over X, increasing the capacitance of SCBs in the network can not fix the issue of network voltage drop.

On the other hand, by comparing the VSI values in Tables 2 and 3, it can be seen that in general, in the first scenario and in the presence of VRs, the voltage stability index has worse conditions than the second scenario (without the presence of VRs). This means that in allocating such equipment, it is very important to consider the voltage stability index as one of the decision criteria. Obviously, if these conditions are not taken into account, a soultion may be chosen that, despite improvements in other parameters, will bring the feeder closer to the voltage collapse condition. For example, a comparison of solutions 1 and 4 in Table 2 indicates that solution 1 is superior to solution 4 in terms of losses and has the same conditions in terms of voltage deviation. But in terms of voltage stability index, Solution 4 is superior to Solution one.

TABLE 4. The φ Values for different scenarios

Scenario No.	MMOPSO		Original MOPSO	
	φ	Solution No.	φ	Solution No.
1	0.2977	4	0.2764	4
2	0.2575	1	0.2437	4

5. CONCLUSION

In this paper, a multi-objective programming model is presented to determine the optimal size and site of VRs and SCBs in long distribution feeders. In this regard, an MMOPSO is employed to solve the proposed model, and the obtained results have been compared with an original MOPSO.

This paper's main aims are voltage level adjustment, loss reduction, voltage profile improvement, voltage stability enhancement, and installation cost minimization. Several important observations can be concluded as follows:

- Utilizing shunt capacitors cannot address the voltage stability margin issues. As a result, the simultaneous implementation of VRs and SCBs is necessary .
- The VSI index of the distribution system faces enormous difficulties by installing VRs, so it is vital to consider the VSI criteria in an optimal allocation of VRs .
- The power losses of the distribution system can be effectively reduced.
- Using proposed MMOPSO leads to more optimal solutions than the original MOPSO.

6. REFERENCES

1. Chakravorty, M. and Das, D., "Voltage stability analysis of radial distribution networks", ***International Journal of Electrical Power & Energy Systems***, Vol. 23, No. 2, (2001), 129-135. DOI: doi.org/10.1016/S0142-0615(00)00040-5
2. Liang, R.-H. and Wang, Y.-S., "Fuzzy-based reactive power and voltage control in a distribution system", ***IEEE Transactions on Power Delivery***, Vol. 18, No. 2, (2003), 610-618. DOI: 10.1109/TPWRD.2003.809740
3. Gu, Z. and Rizy, D.T., "Neural networks for combined control of capacitor banks and voltage regulators in distribution systems", ***IEEE Transactions on Power Delivery***, Vol. 11, No. 4, (1996), 1921-1928. DOI: 10.1109/61.544277
4. Baran, M.E. and Wu, F.F., "Optimal capacitor placement on radial distribution systems", ***IEEE Transactions on Power Delivery***, Vol. 4, No. 1, (1989), 725-734. DOI: 10.1109/61.19265
5. Diab, A.A.Z. and Rezk, H., "Optimal sizing and placement of capacitors in radial distribution systems based on grey wolf, dragonfly and moth–flame optimization algorithms", ***Iranian Journal of Science and Technology, Transactions of Electrical Engineering***, Vol. 43, No. 1, (2019), 77-96. DOI: 10.1007/s40998-018-0071-7
6. Batmani, Y. and Golpîra, H., "Automatic voltage regulator design using a modified adaptive optimal approach", ***International Journal of Electrical Power & Energy Systems***, Vol. 104, (2019), 349-357. DOI: org/10.1016/j.ijepes.2018.07.001
7. Franco, J.F., Rider, M.J., Lavorato, M. and Romero, R., "A mixed-integer lp model for the optimal allocation of voltage regulators and capacitors in radial distribution systems", ***International Journal of Electrical Power & Energy Systems***, Vol. 48, (2013), 123-130. DOI: org/10.1016/j.ijepes.2012.11.027
8. Hock, R.T., Novaes, Y.R.d. and Batschauer, A.L., "A voltage regulator for power quality improvement in low-voltage distribution grids", ***IEEE Transactions on Power Electronics***, Vol. 33, No. 3, (2018), 2050-2060. DOI: 10.1109/TPEL.2017.2693239
9. Siahi, M., Porkar, S., Abbaspour, T.A., Poure, P. and Saadate, S., "Competitive distribution system planning model integration of DG, interruptible load and voltage regulator devices", ***Iranian Journal of Science & Technology, Transaction B: Engineering***, Vol. 34, No. B6, (2010), 619-635. Retrieved from https://www.sid.ir/FileServer/JE/8542010B603.pdf
10. Mohammadi, M., Parsa, M., Alinejad-Beromi, Y. and Moradi, A., "Optimal locating and sizing of unified power quality conditioner-phase angle control for reactive power compensation in radial distribution network with wind generation", ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 2, (2018), 299-306. DOI: 10.5829/ije.2018.31.02b.14
11. El-Ela, A.A.A., El-Sehiemy, R.A. and Abbas, A.S., "Optimal placement and sizing of distributed generation and capacitor banks in distribution systems using water cycle algorithm", ***IEEE Systems Journal***, Vol. 12, No. 4, (2018), 3629-3636. DOI: 10.1109/JSYST.2018.2796847
12. Chamana, M. and Chowdhury, B.H., "Optimal voltage regulation of distribution networks with cascaded voltage regulators in the presence of high pv penetration", ***IEEE Transactions on Sustainable Energy***, Vol. 9, No. 3, (2018), 1427-1436. DOI: 10.1109/TSTE.2017.2788869
13. Efkarpidis, N., Rybel, T.D. and Driesen, J., "Optimal placement and sizing of active in-line voltage regulators in flemish lv distribution grids", ***IEEE Transactions on Industry Applications***, Vol. 52, No. 6, (2016), 4577-4584. DOI: 10.1109/TIA.2016.2599148
14. Safigianni, A.S. and Salis, G.J., "Optimum voltage regulator placement in a radial power distribution network", ***IEEE Transactions on Power Systems***, Vol. 15, No. 2, (2000), 879-886. DOI: 10.1109/59.867188
15. Moradian, S., Jadid, S. and Homaee, O., "Optimal allocation of capacitors with stand-alone var control systems in radial distribution networks", ***International Transactions on Electrical Energy Systems***, Vol. 25, No. 7, (2015), 1333-1348. DOI: org/10.1002/etep.1911
16. Pamshetti, V.B. and Singh, S.P., "Optimal coordination of pv smart inverter and traditional volt-var control devices for energy cost savings and voltage regulation", ***International Transactions on Electrical Energy Systems***, Vol. 29, No. 7, (2019), e12042. DOI: org/10.1002/2050-7038.12042
17. Cao, Y., Wang, Q., Wang, Z., Jermsittiparsert, K. and Shafiee, M., "A new optimized configuration for capacity and operation improvement of cchp system based on developed owl search algorithm", ***Energy Reports***, Vol. 6, (2020), 315-324. DOI: https://doi.org/10.1016/j.egyr.2020.01.010
18. Dixit, M., Kundu, P. and Jariwala, H.R., "Optimal integration of shunt capacitor banks in distribution networks for assessment of techno-economic asset", ***Computers & Electrical Engineering***, Vol. 71, (2018), 331-345. DOI: https://doi.org/10.1016/j.compeleceng.2018.07.014
19. Niknam, T., Ranjbar, A., Ostadi, A. and Shirani, A., "A new approach based on ant algorithm for volt/var control in distribution network considering distributed generation", ***Iranian Journal of Science & Technology, Transaction B, Engineering***, Vol. 29, No. B4, (2005), 385-398. Retrieved from https://www.sid.ir/FileServer/JE/8542005B405.pdf
20. Montazeri, M. and Askarzadeh, A., "Capacitor placement in radial distribution networks based on identification of high potential busses", ***International Transactions on Electrical Energy Systems***, Vol. 29, No. 3, (2019), e2754. DOI: https://doi.org/10.1002/etep.2754
21. Tolba, M.A., Zaki Diab, A.A., Tulsky, V.N. and Abdelaziz, A.Y., "Vlci approach for optimal capacitors allocation in distribution networks based on hybrid PSOGSA optimization algorithm", ***Neural Computing and Applications***, Vol. 31, No. 8, (2019), 3833-3850. DOI: https://doi.org/10.1007/s00521-017-3327-7
22. Kola Sampangi, S. and Thangavelu, J., "Optimal capacitor allocation in distribution networks for minimization of power loss and overall cost using water cycle algorithm and grey wolf optimizer", ***International Transactions on Electrical Energy Systems***, Vol. 30, No. 5, (2020), e12320. DOI: https://doi.org/10.1007/s40998-018-0071-7.
23. Attar, M., Homaee, O., Falaghi, H. and Siano, P., "A novel strategy for optimal placement of locally controlled voltage regulators in traditional distribution systems", ***International Journal of Electrical Power & Energy Systems***, Vol. 96, (2018), 11-22. DOI: https://doi.org/10.1016/j.ijepes.2017.09.028
24. Szuvovivski, I., Fernandes, T.S.P. and Aoki, A.R., "Simultaneous allocation of capacitors and voltage regulators at distribution networks using genetic algorithms and optimal power flow", ***International Journal of Electrical Power & Energy Systems***, Vol. 40, No. 1, (2012), 62-69. DOI: https://doi.org/10.1016/j.ijepes.2012.02.006
25. Das, S., Das, D. and Patra, A., "Operation of distribution network with optimal placement and sizing of dispatchable dgs and shunt capacitors", ***Renewable and Sustainable Energy Reviews***, Vol. 113, (2019), 109219. DOI: https://doi.org/10.1016/j.rser.2019.06.026
26. Elmitwally, A. and Eldesouky, A., "An approach for placement and sizing of capacitor banks in distribution networks with distributed wind generation", ***International Transactions on***

Electrical Energy Systems, Vol. 23, No. 4, (2013), 539-552. DOI: https://doi.org/10.1002/etep.1611

27. Haghighat, H., "Energy loss reduction by optimal distributed generation allocation in distribution systems", ***International Transactions on Electrical Energy Systems***, Vol. 25, No. 9, (2015), 1673-1684. DOI: https://doi.org/10.1002/etep.1907
28. Elsayed, A.M., Mishref, M.M. and Farrag, S.M., "Distribution system performance enhancement (egyptian distribution system real case study)", ***International Transactions on Electrical Energy Systems***, Vol. 28, No. 6, (2018), e2545. DOI: https://doi.org/10.1002/etep.2545
29. Mishra, A., Farzinfar, M., Bahadornejad, M. and Nair, N.C., "Evaluating the impact of different pv control strategies on distribution network operation", In 2014 Australasian Universities Power Engineering Conference (AUPEC), (2014), 1-6. DOI: 10.1109/AUPEC.2014.6966591
30. Farzinfar, M., Shafiee, M. and Kia, A., "Determination of optimal allocation and penetration level of distributed energy resources considering short circuit currents", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 33, No. 3, (2020), 427-438. DOI: 10.5829/IJE.2020.33.03C.07
31. Aolaritei, L., Bolognani, S. and Dörfler, F., "Hierarchical and distributed monitoring of voltage stability in distribution networks", ***IEEE Transactions on Power Systems***, Vol. 33, No. 6, (2018), 6705-6714. DOI: 10.1109/TPWRS.2018.2850448
32. Jasmon, G.B. and Lee, L.H.C.C., "New contingency ranking technique incorporating a voltage stability criterion", ***IEE Proceedings C - Generation, Transmission and Distribution***, Vol. 140, No. 2, (1993), 87-90. DOI: 10.1049/ip-c.1993.0012
33. Gubina, F. and Strmcnik, B., "A simple approach to voltage stability assessment in radial networks", ***IEEE Transactions on Power Systems***, Vol. 12, No. 3, (1997), 1121-1128. DOI: 10.1109/59.630451
34. Tolba, M.A., Diab, A.A.Z., Tulsky, V.N. and Abdelaziz, A.Y., "Vlci approach for optimal capacitors allocation in distribution networks based on hybrid PSOGSA optimization algorithm", ***Neural Computing and Applications***, Vol. 31, No. 8, (2019), 3833-3850. DOI: https://doi.org/10.1007/s00521-017-3327-7
35. de Oliveira, L.W., Carneiro Jr, S., De Oliveira, E.J., Pereira, J., Silva Jr, I.C. and Costa, J.S., "Optimal reconfiguration and capacitor allocation in radial distribution systems for energy losses minimization", ***International Journal of Electrical Power & Energy Systems***, Vol. 32, No. 8, (2010), 840-848. DOI: https://doi.org/10.1016/j.ijepes.2010.01.030
36. Haghifam, M.-R. and Malik, O., "Genetic algorithm-based approach for fixed and switchable capacitors placement in distribution systems with uncertainty and time varying loads", ***IET Generation, Transmission & Distribution***, Vol. 1, No. 2, (2007), 244-252. DOI: 10.1049/iet-gtd:20045267
37. Ali, E.S., Abd Elazim, S.M. and Abdelaziz, A.Y., "Improved harmony algorithm for optimal locations and sizing of capacitors in radial distribution systems", ***International Journal of Electrical Power & Energy Systems***, Vol. 79, (2016), 275-284. DOI: https://doi.org/10.1016/j.ijepes.2016.01.015
38. de Araujo, L.R., Penido, D.R.R., Carneiro, S. and Pereira, J.L.R., "Optimal unbalanced capacitor placement in distribution systems for voltage control and energy losses minimization", ***Electric Power Systems Research***, Vol. 154, (2018), 110-121. DOI: https://doi.org/10.1016/j.epsr.2017.08.012
39. Nazarpour, D. and Sattarpour, T., "Assessing the impact of size and site of dgs and sms in active distribution networks for energy losses cost", ***International Journal of Engineering, Transactions A: Basics***, Vol. 28, No. 7, (2015), 1002-1010. DOI: 10.5829/idosi.ije.2015.28.07a.06
40. Kennedy, J. and Eberhart, R., "Particle swarm optimization", In Proceedings of ICNN'95 - International Conference on Neural Networks. Vol. 4, (1995), 1942-1948. DOI: 10.1109/ICNN.1995.488968
41. Coello, C.A.C., Lamont, G.B. and Van Veldhuizen, D.A., "Evolutionary algorithms for solving multi-objective problems, Springer, Vol. 5, (2007). Retrieved from https://link.springer.com/content/pdf/10.1007/978-0-387-36797-2.pdf
42. Coello, C.A.C., Pulido, G.T. and Lechuga, M.S., "Handling multiple objectives with particle swarm optimization", ***IEEE Transactions on Evolutionary Computation***, Vol. 8, No. 3, (2004), 256-279. DOI: 10.1109/TEVC.2004.826067
43. Moravej, Z., Jazaeri, M. and Gholamzadeh, M., "Optimal coordination of distance and over-current relays in series compensated systems based on MAPSO", ***Energy Conversion and Management***, Vol. 56, (2012), 140-151. DOI: https://doi.org/10.1016/j.enconman.2011.11.024
44. El-Hawary M E, M.G.A.N., "Stochastic optimal load flow using newton raphson iterative technique", ***Electric Machines Power Systems***, Vol. 15, (1988), 371-380. DOI: https://doi.org/10.1080/07313568808909347
45. Carlisle, A. and Dozier, G., "An off-the-shelf PSO" In Proceedings of the Workshop on Particle Swarm Optimization, Indianapolis, USA, (2001), 1-6. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.589.485&rep=rep1&type=pdf

Persian Abstract

چکیده

نصب بانک‌های خازنی و رگولاتورهای ولتاژ در شبکه‌های توزیع برق از جمله راهکارهای موثر در کنترل توان راکتیو به منظور بهبود پروفیل ولتاژ فیدرها و کاهش تلفات شبکه به حساب می‌آید. با این حال، بهره‌برداری از شبکه‌های توزیع در حضور رگولاتورهای ولتاژ ممکن است منجر به عدول پارامترهای پایداری ولتاژ از مقادیر مجاز و استاندارد گردد. به منظور رفع این مشکل، در این مقاله یک مدل جدید بر مبنای یک مسئله بهینه سازی چندهدفه پیشنهاد شده است که قادر به مکان‌یابی همزمان بهینه بانک‌های خازنی و رگولاتورهای ولتاژ با هدف بهبود پروفیل ولتاژ، کاهش تلفات و همچنین کاهش هزینه‌ها می‌باشد. در مدل پیشنهادی، از یک شاخص پایداری ولتاژ به منظور جلوگیری از ایجاد ناپایداری ولتاژ در شبکه توزیع هنگام جایابی ادوات کنترلی استفاده شده است. همچنین، برای حل مدل پیشنهادی یک الگوریتم PSO اصلاح شده، مبتنی بر وزن دهی دینامیکی ذرات و استفاده از عملگر جهش ارایه شده است. در ادامه به منظور انتخاب بهترین راه‌حل‌ها از مجموعه جواب‌های پرتو، الگوریتم استنتاج فازی بکار گرفته شده است. به منظور بررسی کارآیی الگوریتم پیشنهادی، یک سیستم تست واقعی مورد استفاده قرار گرفته است. تصدیق کارایی الگوریتم MMOPSO پیشنهادی با انجام مقایسه با جواب‌های بهینه حاصل از الگوریتم MOPSO انجام شده است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Bandwidth Management with Congestion Control Approach and Fuzzy Logic

G. Azami, H. Gholizade Narm*

Faculty of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran

P A P E R I N F O

Paper history:
Received 12 November 202
Received in revised form 07 February 2021
Accepted 15 February 2021

Keywords:
Bandwidth Management
Congestion Control
Fuzzy PID Controller
TCP/IP Networks

A B S T R A C T

One of the problems with today's TCP/IP networks is their transmission system. If the bandwidth of a network is full, human and physical factors must be used for a new transmission system with a higher capacity to provide its bandwidth, which is very time consuming and costly. In this article, we proposed a method that in addition to the optimal use of available bandwidth, if the network capacity is full, it will be automatically transferred to a higher bandwidth network. For this purpose first, by designing a fuzzy PID controller for the existing network, it is tried to congestion control them and make use of it. It can be seen that the proposed controller performs much better in terms of an output response, following the queue length, stability and uncertainly, compared to the classical controller. If the input data to the network is increased, more packets are lost and this reduces the quality of the network. To solve this problem by using bandwidth management, by considering the threshold for packets loss in each network, if exceeding this limit, the existing network is switched to a network with a higher capacity and the problem of bandwidth and network quality is solved and causes subscriber satisfaction.

doi: *10.5829/ije.2021.34.04a.15*

NOMENCLATURE

W	Average TCP window size (packet)	e	Error
q	Average queue length (packet)	$\dot{e}$	Derivative of error
R(t)	Packets return time (s)	μ	Membership function
C	Link capacity (packet/s)	Bw	Bandwidth
Tp	propagation delay (s)	NB	Negative big
N	Number of resources	NM	Negative medium
P(t)	probability of the packet being marked	NS	Negative small
K_p	Proportional gain	ZO	Zero
K_i	Integral gain	PB	Positive big
K_d	Derivative gain	PM	Positive medium
K_u	Gain (Nichols method)	PS	Positive small
T_u	Period (Nichols method)	RTT	Round trip time(s)

INTRODUCTION

The subject of data congestion in networks based on Transmission Control Protocol (TCP) has become a major research topic in recent years [1]. Floyd and Jacobson [2] offered two solutions to the problem of congestion. The first solution suggests the solution in the initial and final points of sending packets, i.e. sender and receiver, and the second solution provides the adoption of the policy in the middle points of the network, i.e. routers. The first solution is generally interpreted as congestion control through TCP and the second solution is interpreted as Active Queue Management (AQM). In 1993, the Random Early Detection (RED) method was

*Corresponding Author Institutional Email: *gholizade@shahroodut.ac.ir* (H. Gholizade Narm)

Please cite this article as: G. Azami, H. Gholizade narm, Bandwidth management with congestion control approach and fuzzy logic, International Journal of Engineering Transactions A: Basics Vol. 34, No. 04, (2021) 891-900

projected [2], according to which the middle points of the network are randomly deleted before the packing queue. Through this method, it was determined according to the studies that it does not have the required efficiency and cause of the fluctuation of the queue length [3]. Consequently, numerous methods have been projected to develop the projected method, including Stability RED, Fuzzy RED [4], Adaptive RED [5], PD-RED (Proportional-Integral) methods. Notably, some have also tried to offer newer methods for active queue control, including BLUE [6], Multi-GREEN [7]. The nonlinear RED model has also been applied through the TRED method to develop the RED method [8]. Smart methods are also used for congestion control. A new AQM controller called a feed-forward neural network [9] has been projected to effectively control network congestion through establishing queue lengths. This pattern exhibits nonlinear and dynamic network traffic and predicts the future value of queue length. Moreover, we have created a good interaction between network operational capacity and delay through a learning system with loss prediction (LP) and reinforcement learning (RL) [10]. In another study, the problem of density in Bandwidth Delay Product (BDP) networks was solved through presenting a new Elastic-TCP method, resulting in a higher power than other Media Transport Protocol (MTP) [11].

Through mathematical models, it is feasible to envisage the behaviour of the system in different modes and conditions and to adopt the essential policy in dealing with such behaviour. The TCP network has always been regarded by researchers as a complex nonlinear system whose dynamics are frequently changing. A good deal of effort has been completed to mathematically recognize computer network models [12]. This is while one of the best models in this field can be regarded as the Misra model, introduced in 2000 [13]. This model is based on the fluid flow model and provides a nonlinear model of the network. Through presenting this model, different individuals attempted to use control theory to provide a method for active queue control. The basic idea was to use a control rule to accidentally remove the packets in the queue before the queue fills out.

The first controller to be offered was a simple Proportional Integral (PI) controller [14]. The PID controller was introduced, which performed better than the PI and RED controllers, and the process of changing parameters has less oscillation, leading to better queue adjustment. In this method, it was seen that increasing resources and the number of routers is better than control. PI and RED actuators operate and its oscillation is much less than the above methods, but in response to turbulence and uncertainty, the response is slower. A nonlinear congestion control model was used to overcome this problem; an example of which is the use of a slip mode controller [15]. In this method and through finding a rule of feedback control, the model is such that the source is stable and its application to the system has significantly overcome the turbulence and uncertainty. In another study, it uses PID-derived feedback control to correct AQM limitations, monitoring dynamic network changes at any time, and controlling congestion in these environments. In this method, it is observed that dynamically, tracking the length of the queue, the rate of loss of the package, and the use of the link are significantly improved [16].

Much research has been done on fuzzy logic about congestion control. One of the methods is presented in literature [17], which uses fuzzy logic to solve the problem of linearization and parameterization of RED and has developed the RED method with this device. To solve the problem of congestion in wireless sensor networks used in medical, agricultural and military equipment, a method has been proposed [18] in which, using a combination of fuzzy logic and PID, congestion it is detected earlier, and the length of the buffer queue is adjusted by determining the probability of packet loss. In another study, a fuzzy sliding mode congestion control (FSMC) algorithm is presented in literature [19] which, the queue buffer length is adjusted in congestion nodes through a new model of congestion control between the transfer layer and the MAC layer, leading to a decrease in delay and the packet loss and also, to improvement in network power.

In this study, we first examine the issue of active queue AQM management as the basis for different control systems and then obtain the mathematical and dynamic model of the TCP network. In the following, we examine two issues simultaneously. First, the Fuzzy PID control method is presented automatically to adjust the PID parameters, so that in addition to setting the parameters, we can automatically improve the disturbance and uncertainty, and compare it with the classical control method, and point out its advantages. Then, we control the congestion in the network according to the design. If the congestion control is no longer effective, we need to design a mechanism that allows us to use a network with more bandwidth and better performance with Bandwidth management. We propose a new method to solve this problem. By designing three networks with different bandwidths and managing this network, in addition to the optimal use of each network through congestion control, if congestion in each network significantly increases, by taking into account the threshold and using switching system. We moved the existing network to a network with more bandwidth. In this way, the appropriate bandwidth is allocated to each network. The important advantage of this method is that it always allocates the appropriate bandwidth to the subscribers and makes them satisfied.

2. ACTIVE QUEUE MANAGEMENT

Active queue management is applied in the middle points and to control congestion. The initial encounters with the problem of congestion in the middle points were such that the router would insert the packets into the buffer for processing, and if the buffer was full, the newly received packets would be destroyed. In this case, the packets that reached the buffer were processed earlier, which is the same as the implementation of the First-In-First-Out (FIFO) method. This method is called Drop Tail. The instability of the queue in the middle points causes the difference between delays, and this is inappropriate for Real-Time users. In AQM, the queue length (number of packets left in a queue) is tried to remain at a reference point determined by the designer or operator toward preventing packet overflows. Any deviation of the queue length from the reference results in changing from the router to the network resources through sending commands so that the queue length returns to the reference point. AQM can be applied as a controller for the TCP network to keep the queue length constant, assuming that its output is the probability of destroying packets and its input is the amount of congestion. Therefore, the theory of control can be applied to form a control law, which is the same as the possibility of eliminating packets. Many methods based on this principle have been projected for AQM, such as RED, PI, PID, etc [20].

2. 1. TCP/IP Network Mathematical Model

A mathematical model of the network was introduced in 2000. Based on the studies, this model is closer to the real system. It describes the mathematical model of an AQM that supports TCP currents and designs an AQM-resistant controller. The controller is designed for the N data flow of TCP. The shutdown time (RTT) of each stream is defined as follows:

$$R(t) = \frac{q(t)}{c} + T_p \tag{1}$$

One dynamic model of TCP behaviour is the use of fluid flow and statistical differential equation analysis [13]. These dynamics include the same N number of TCP sources and a router to determine the average amount of changes and dynamics queue the router (Figure 1).

This model depends on key network changes and it is described through the following nonlinear differential equations.

$$\dot{W}(t) = \frac{1}{R(t)} - \frac{w(t)}{2}\frac{W(t-R(t))}{R(t-R(t))}P(t-R(t-R(t))) \tag{2}$$

The above equation consists of two increasing and decreasing sections. The increasing section indicates that the amount of compression window increases by one unit for the time changes of a time shift. This happens during the TCP compression prevention phase. The decreasing section indicates that the value of the window decreases to one half in the event of a loss. The second part of the AQM model is for the middle points, which should be written based on the dynamics of queue length changes. The dynamics of this section are expressed as Equation (3).

$$\dot{q} = \begin{cases} -C + \frac{N(t)}{R(t)}W(t) & q \succ 0 \\ \max\left\{0, -C + \frac{N(t)}{R(t)}W(t)\right\} & q = 0 \end{cases} \tag{3}$$

Equation (3) includes two increasing and decreasing sections. The decreasing section, as the capacity of the midpoints, represents the amount of packet processing per time unit, and the increasing section represents the number of packets reaching queue per time unit. In this equation, the queue length is modelled through the difference between NW/R and the C link capacity. In Equations (1), (2), and (3), the parameters are defined as follows:

W: Average TCP window size (packets); q: Average queue length (packets); R(t): packets return time (seconds); C: Link capacity (seconds/pack); q(t)/C: queue delay; Tp: propagation delay (seconds); P(t): the probability of the packet being marked in the router. By aligning the nonlinear Equations of (1), (2), and (3) around the working point and assuming that N(t)= N and C(t)= C and placing $\dot{w} = 0$ and $\dot{q} = 0$ we reach the following values.

$$\begin{aligned} \dot{w} &= 0 \Rightarrow w_0^2 p_0 = 2 \\ \dot{q} &= 0 \Rightarrow w_0 = \frac{R_0 C}{N} \\ R_0 &= \frac{\dot{q}}{C} + T_p \end{aligned} \tag{4}$$

In linearization, we do not consider the dependence of $t - R$ time delay parameters on the length of the queue (q) and assume that the time delay is replaced through a constant value $t - R_0$. In other words, we do not consider the dependence of the return time depending on the queue length in the dynamic parameters.

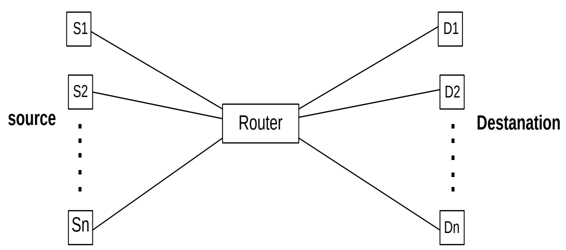


Figure 1. Network topology with N similar sources TCP and a router

$$\dot{w}(t)=\frac{1}{R_0}-\frac{w(t)}{2}\frac{w(t-R_0)}{\frac{q(t-R_0)}{C}+T_p}p(t-R_0)$$

$$\dot{q}(t)=\begin{cases} -c+\dfrac{N}{\frac{q(t-R_0)}{C}}w(t) & q\succ 0 \\ \max\left\{0,-c+\dfrac{N}{\frac{q(t-R_0)}{C}}w(t)\right\} & q=0\end{cases} \qquad (5)$$

Therefore, linearization at the equilibrium point leads to the following equations:

$$\begin{aligned}\delta\dot{w}(t)&=-\frac{2N}{R_0^{\,2}C}\delta w(t)-\frac{R_0C^2}{2N^2}\delta p(t-R_0)\\ \delta\dot{q}(t)&=\frac{N}{R_0}\delta w(t)-\frac{1}{R_0}\delta q(t)\end{aligned} \qquad (6)$$

where,

$$\delta w=w-w_0,\ \delta q=q-q_0,\ \delta p=p-p_0 \qquad (7)$$

From the Laplace transform of the Equations (6) and (7), the system conversion function is obtained as follows. the nonlinear Equations (1), (2), and (3) around the operating point (w_0, q_0, p_0) and assuming that N(t)=N, C(t)=C, and by putting $\dot{q}=0$, $\dot{w}=0$, we get the dynamic Equation (4).

$$\frac{\frac{(RC)^3}{4N^2}e^{-Rs}}{(Rs+1)(\frac{R^2C}{2N}s+1)} \qquad (8)$$

This dynamic equation of congestion control is in the TCP/IP network, as the basis for our further discussion.

3. CONGESTION CONTROL USING FUZZY PID

Due to the change of network parameters, the need for a controller that can respond to changes properly and keep the system stable seems necessary. One of the suitable methods for this work is fuzzy methods that can adapt the system to changes and uncertainties and also adjust the parameters. In congestion control, fuzzy methods have been able to receive changes in the network and adjust the controller parameters based on it and send the appropriate control signal to achieve the desired output [21]. In this research, the fuzzy controller method has been used to control congestion. In this method, using the swarm dynamics, we design a PID controller so that the controller details are adjusted by the above algorithm and then a suitable control signal is generated for the process.

In the fuzzy control method, using the error and error derivative and applying the appropriate rules on these parameters, the PID controller coefficients are adjusted so that a suitable signal is generated for the process. This is shown in Figure 2 [22].

3. 1. Mathematical Topics To implement the fuzzy PID control, we define the control signal U (t) based on the PID as follows.

$$\begin{aligned}G_C(S)&=K_p+\frac{K_i}{S}+K_DS=K_p(1+\frac{1}{T_iS}+T_dS)\\ K_i&=\frac{K_p}{T_i},K_d=K_pT_d\end{aligned} \qquad (9)$$

To determine the coefficients of the PID controller, we can use the Ziegler and Nichols method. For this purpose, we put the controller in P mode, and by increasing the controller gain, we act so that the system output is on the verge of fluctuation. In this case, the gain is K_u and Consider the oscillation period T_u. The following values of PID controller values are obtained as follows [23]:

$$K_p=0.6K_u,K_i=2k_p/k_u,K_d=k_pT_u/8 \qquad (10)$$

After obtaining the controller finances, which we also use in the fuzzy controller, we describe the fuzzy system in this research. To begin with, instead of K_p, K_d and K_i, we get K_p', K_d' and α, where $\alpha=T_i/T_d$ is. As a result, the value of K_i is obtained using Equation (9) and the value of α is obtained as follows.

$$K_i=\frac{K_p}{T_i},T_i=\alpha T_d,T_d=\frac{K_d}{K_p}\rightarrow K_i=\frac{K_p^{\,2}}{\alpha K_d} \qquad (11)$$

In the following, instead of obtaining the absolute values of the controlling fines, we obtain their normalized values by the following method.

$$\begin{aligned}k_p'&=\frac{k_p-k_p^{\min}}{k_p^{\max}-k_p^{\min}}\in[0,1]\\ k_d'&=\frac{k_d-k_d^{\min}}{k_d^{\max}-k_d^{\min}}\in[0,1]\end{aligned} \qquad (12)$$

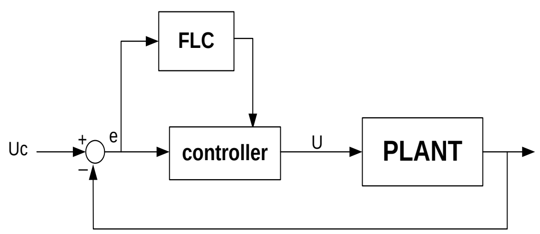


Figure 2. Fuzzy controller block diagram network topology with N similar sources TCP and a router

According to Figure 3 with error input and error derivative, fuzzy system outputs are defined as follows. The values of K_p, K_d and K_i are calculated for the values of K_p', K_d' and α To obtain fuzzy PID coefficients, we must carefully define fuzzy rules. For this purpose, we define the rules as follows:

$$\begin{aligned} &If\ e\ is\ A_i\ and\ \dot{e}\ is\ B_i\ ,then\ k'_p\ is \begin{Bmatrix} Big \\ Small \end{Bmatrix} \\ &k'_d\ is \begin{Bmatrix} Big \\ Small \end{Bmatrix} \\ &\alpha = 2,3,4,5 \end{aligned} \tag{13}$$

For this purpose, we specify the membership function for e and $\dot{e}$ in Figure 4 [21, 24].

For get K_p' and K_d' we consider the membership functions graphically as shown in Figure 5. According to Figure 5, we consider the membership function as follows [25]:

$$\begin{aligned} &\mu_{Small}(x) = \min(-\frac{1}{4}\ln x, 1) \\ &\mu_{Big}(x) = \min(-\frac{1}{4}\ln(1-x), 1) \end{aligned} \tag{14}$$

We also use a fuzzy singleton for α.

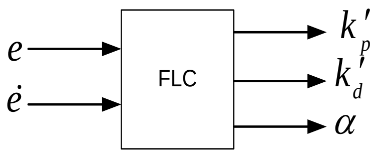


Figure 3. Fuzzy system input and output

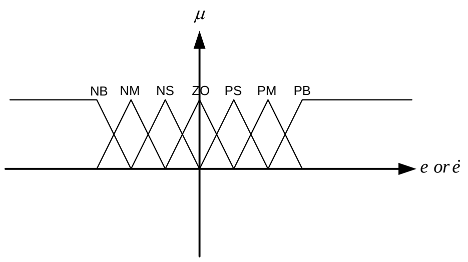


Figure 4. Membership functions for e and $\dot{e}$

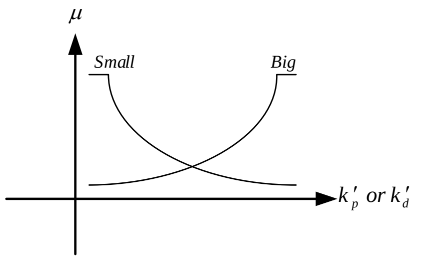


Figure 5. Membership functions for K_p' and K_d'

In the following, we consider the following conditions to determine the fuzzy rules.

$$\begin{aligned} &Rule\ i: \\ &If\ e\ is\ A_i\ ,\dot{e}\ is\ B_i\ ,Then\ K'_p\ is\ C_i\ ,K'_d\ is\ D_i\ ,\alpha\ is\ \alpha_i \\ &\mu_i \propto \mu_{A_i}(e)\mu_{B_i}(\dot{e}) \quad \sum \mu_i = 1 \end{aligned} \tag{15}$$

As a result, the output values of the fuzzy system are equal to:

$$K'_p = \sum_i \mu_i k'_{p_i} \ ,K'_d = \sum_i \mu_i k'_{d_i} \ ,\alpha = \sum_i \mu_i \alpha_i \tag{16}$$

To obtain the fuzzy system rules table, we use the appropriate answer. In this case, for different points of this answer, we must analyze the performance of the fuzzy system to obtain its rules. Suppose the response of a system is in Figure 6.

The purpose of fuzzy system design is to achieve the appropriate response of Figure 6. For this purpose, we follow the answer path. If we want to get from point a_1 to b_1, we have to increase K_p K_i and decrease K_d. Also, if we want to change the answer around the reference value, to follow the path from b_1 to c_1, we have to decrease K_p K_i and increase K_d. Then, in order to reach from c_1 to d_1, K_p and K_i have to increase and K_d decrease. Also, similar to the previous case, to reach from point a_1 to b_1, must be increased K_p', K_d' and α must be reduced by using relations $K_i = \frac{K_p}{T_i}$ and $\alpha = T_i/T_d$. Other changes are obtained according to the answer.

For this purpose, using the above method and the membership function of Figures 4 and 5 we obtain the fuzzy rules, which you can see in the Tables 1, 2 and 3 [21, 22, 26].

In Table 1, we consider the gain changes K_p' as follows. If the error ***e(k)*** is a large value (NB), regardless of the value of the error derivative ***ė(k)*** , we choose a large gain of K_p' (B). If the error value decreases (NM), and also the error derivative value decreases (NM or NS), we consider the value of gain K_p' to be large(B). If the error value is zero (ZO) and the error derivative value is

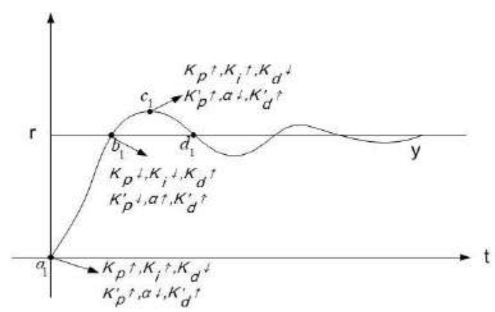


Figure 6. The response of the second-order system for determining fuzzy rules

low, we consider the value gain K_p' to be small(S), unless the derivative of the error is zero, in this case, we choose a large gain of K_p' (B). In the same way, different amounts of gain K_d' and α obtained based on the error and the error derivative.

From the rules of Tables 1 to 3, the values of K_p', K_p' obtained from Equations (11) and (12), and PID control coefficients are obtained as follows:

$$\begin{aligned} K_P &= K_p^{\min} + (k_p^{\max} - k_p^{\min})k'_p \\ K_d &= K_d^{\min} + (k_d^{\max} - k_d^{\min})k'_d \\ K_i &= \frac{K_p^2}{\alpha K_d} \end{aligned} \tag{17}$$

where in:

$$\begin{cases} k_p^{\min} = 0.32k_u \\ k_p^{\max} = 0.6k_u \end{cases} \quad \begin{cases} k_d^{\min} = 0.08k_uT_u \\ k_d^{\max} = 0.15k_uT \end{cases} \tag{18}$$

Using the Zicklor and Nichols method [27, 28], we can first adjust the PID parameters and use the values k and t obtained from it to design the rules of the fuzzy system and extract the final PID values to produce the fuzzy controller control signal. We will explain this in the simulation section of this article.

TABLE 1. Fuzzy system rules for K_p'

		$\dot{e}(k)$						
		NB	**NM**	**NS**	**ZO**	**PS**	**PM**	**PB**
	NB	B	B	B	B	B	B	B
	NM	S	B	B	B	B	B	S
e(k)	**NS**	S	S	B	B	B	S	S
	ZO	S	S	S	B	S	S	S
	PS	S	S	B	B	B	S	S
	PM	S	B	B	B	B	B	S
	PB	B	B	B	B	B	B	B

TABLE 2. Fuzzy system rules for K_d'

		$\dot{e}(k)$						
		NB	**NM**	**NS**	**ZO**	**PS**	**PM**	**PB**
	NB	S	S	S	S	S	S	S
	NM	B	B	S	S	S	B	B
e(k)	**NS**	B	B	B	S	B	B	B
	ZO	B	B	B	B	B	B	B
	PS	B	B	B	S	B	B	B
	PM	B	B	S	S	S	B	B
	PB	S	S	S	S	S	S	S

TABLE 3. Fuzzy system rules for α

		$\dot{e}(k)$						
		NB	**NM**	**NS**	**ZO**	**PS**	**PM**	**PB**
	NB	2	2	2	2	2	2	2
	NM	3	3	2	2	2	3	3
e(k)	**NS**	4	3	3	2	3	3	4
	ZO	5	4	3	3	3	4	5
	PS	4	3	3	2	3	3	4
	PM	3	3	2	2	2	3	3
	PB	2	2	2	2	2	2	2

4. BANDWIDTH MANAGEMENT WITH CONGESTION CONTROL APPROACH AND FUZZY LOGIC

In previous discussions, it was explained how to control network congestion with a fuzzy method so that the network can remain stable; but in some cases, the network inputs are so high that the existing network no longer responds to network data despite the congestion control system. In this case, the data packets are largely deleted and the congestion control response to the system is not available. An appropriate solution must be found for this issue. In this paper, the proposed solution is to set a threshold for deleting packages. As long as the removal of packets is less than this threshold, the existing system will work without change. If the amount of packet deletion exceeds the threshold, use a switching system to transfer the network to a network with more bandwidth so that it can have more capacity to transfer data. With this method, in addition to optimizing the use of the existing network, if the network is full, we can transfer to the network with more capacity using the designed system and keep the network stable and cause more satisfaction of data subscribers. We display this in Figure 7.

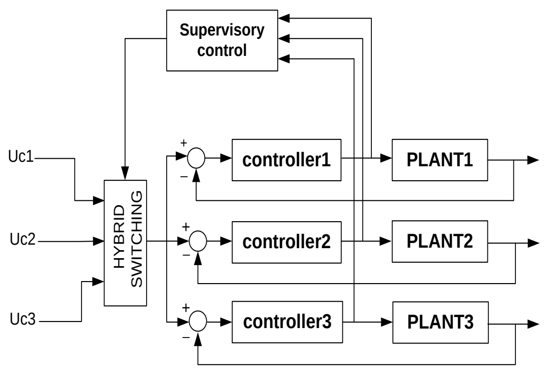


Figure 7. Bandwidth management

5. SIMULATION

To simulate bandwidth management with congestion control and fuzzy logic approach, we consider three networks with bandwidths of 10Mb/s, 50Mb/s, and 100Mb/s. We first test the network with the specifications of N=60, RTT=200ms, and C=152paket/s or bandwidth of 10mb/s. With these features, the network's congestion dynamics are equal to:

$$\frac{194.1e^{-0.2t}}{S^2+24.67S+98.3} \tag{19}$$

In this mode, we simulate congestion control for both classical control modes and fuzzy logic for the buffer with the desired queue of 100 using MATLAB software.

As you can see in Figure 8, the fuzzy PID controller outperforms the classical PID controller in terms of output response and system speed in terms of congestion control.

In Figure 9, you can see that the control signal U shows a value of 50, which is the same amount of packet loss in this network, and its value is acceptable. Next, we increase the input data to the network from N=60 to N=100, and we want to see how congestion control works with existing network features. In this case, the network dynamics with N=100, RTT=200ms, and C=152packet/s or bandwidth of 10mb/s is obtained as follows:

$$\frac{116e^{-0.2t}}{S^2+37.77S+163.8} \tag{20}$$

You can see the simulation of the above network with the changes made with the desired queue of 100 for the output response and the control signal in Figures 10 and 11.

According to Figure 10, it can be seen that by increasing the input data to the network from N = 60 to N = 100, you follow the desired value of 100 well, but the control signal, which is the packets loss, increases from 50 to 141, which packets loss is much higher than the standard value and is not acceptable for the existing network.

To solve this problem, we consider a network with a higher capacity and a threshold value of 80 to remove packets. If the control signal, which is the removal of packets, exceeds the threshold value of 80, the existing network does not respond to the input data and must be switched to a higher capacity network. In this case, a network with N=100 and RTT=200ms and capacity

Figure 8. Congestion Control Output Based on Controller Fuzzy PID and Classic PID with BW=10mb/s, N=60, Queue Length=100

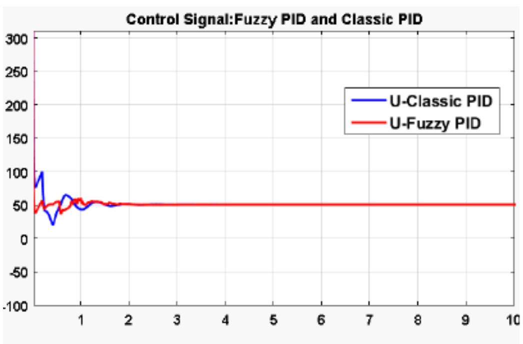


Figure 9. Control Signal Based on Controller Fuzzy PID and Classic PID with BW=10mb/s, N=60, Queue Length=100

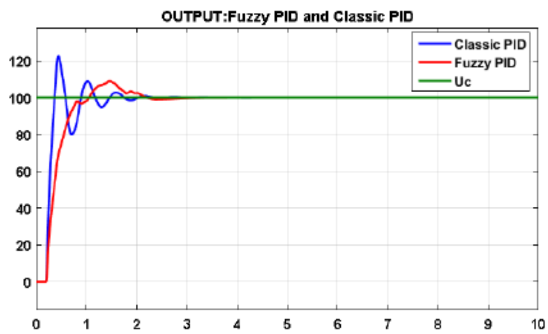


Figure 10. Congestion Control Output Based on Controller Fuzzy PID and Classic PID with BW=10mb/s, N=100, Queue Length=100

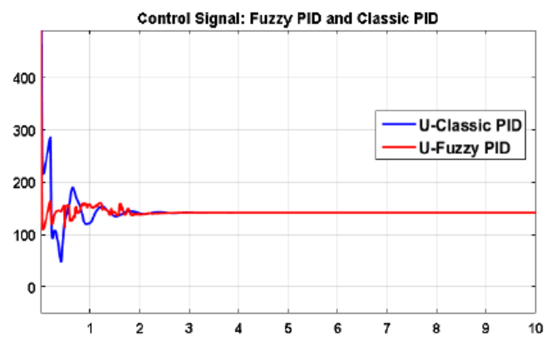


Figure 11. Control Signal Based on Controller Fuzzy PID and Classic PID with BW=10mb/s, N=100, Queue Length=100

consider C=50mb/s. In this case, the network dynamics are as follows:

$$\frac{88e^{-0.2t}}{S^2+11.553S+32.765} \tag{21}$$

We performed the simulation for this network with the desired queue value of 100.

As observed in Figures 12 and 13, in the new network, in addition to controlling congestion and making optimal use of it, the queue length of 100 is well followed and the packets' loss is greatly decreased (41 packets). If we would apply the previous network with this number of resources, the number of packets loss would reach 141, which is not acceptable.

5. 1. Effects of Uncertainty in the Fuzzy PID Controller To study the effects of uncertainty in the fuzzy PID controller designed in Section 3.1, we first change one of the plant coefficients and observe the system response to this change. To change, we assume that the f coefficient b changes as follows.

$$b+[\alpha^-,\alpha^+] \tag{22}$$

We considered the value of α equal to 0.5. In this case, according to the value of b in the plant (19), which is 24.67, we apply the values of b as follows.

$$[24.17,\ 25.17] \tag{23}$$

First, we consider the value b to be 24.17 and apply it to plant (19). The plant is obtained as follows.

$$\frac{194.1e^{-0.2t}}{S^2+24.17S+98.3} \tag{24}$$

We perform a simulation with these specifications.

Then, considering b equal to 25.17, the plant it is obtained as follows.

$$\frac{194.1e^{-0.2t}}{S^2+25.17S+98.3} \tag{25}$$

The above plant simulation is obtained as follows.

We simulated the same conditions for the plant (21) with a coefficient (b=11.553) with a value of α equal to 0.5. The value of b is obtained as follows:

$$[11.053,\ 12.053] \tag{26}$$

By applying the value b to the plant, the following values are obtained and we perform a simulation for each.

$$\frac{88e^{-0.2t}}{S^2+11.053S+32.765},\ \frac{88e^{-0.2t}}{S^2+12.053S+32.765} \tag{27}$$

As shown in Figures 14 to 17, when changing the coefficient b, the fuzzy PID controller does not change,

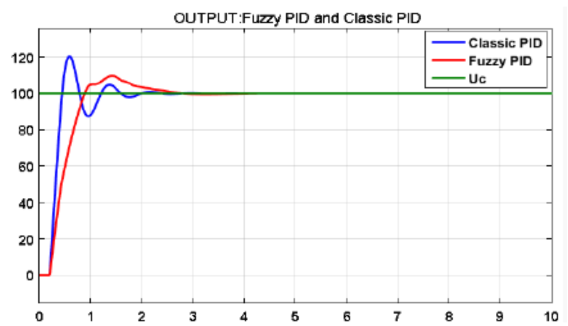


Figure 12. Congestion Control Output Based on Controller Fuzzy PID and Classic PID with BW=50mb/s, N=100, Queue Length=100

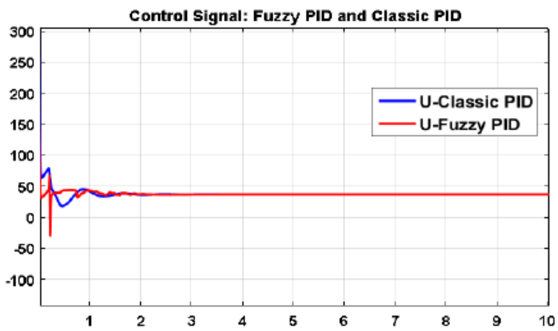


Figure 13. Control Signal Based on Controller Fuzzy PID and Classic PID with BW=50mb/s, N=100, Queue Length=100

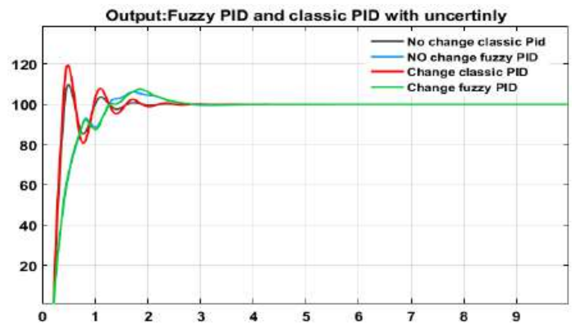


Figure 14. The output of classical PID and fuzzy PID control systems by changing the coefficient b (b=24.17)

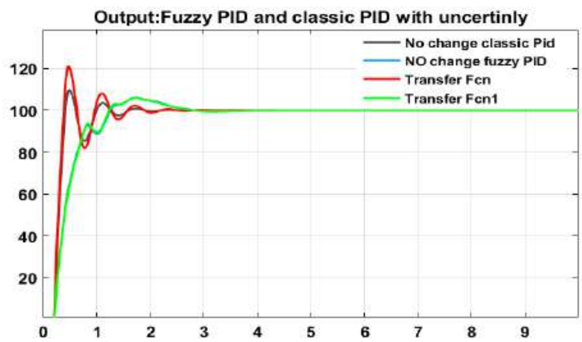


Figure 15. The output of classical PID and fuzzy PID control systems by changing the coefficient b (b=25.17)

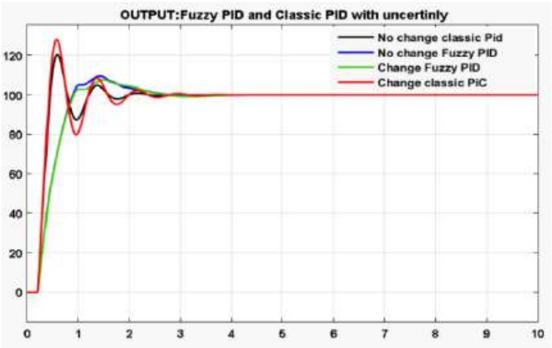


Figure 16. The output of classical PID and fuzzy PID control systems by changing the coefficient b (b=11.053)

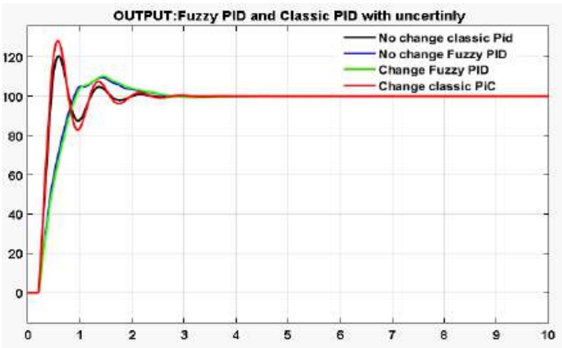


Figure 17. The output of classical PID and fuzzy PID control systems by changing the coefficient b (b=12.053)

but the classic PID controller significantly changes. In this case, we conclude that the fuzzy PID controller responds well to uncertainty.

6. CONCLUSION

In this study, we examined two issues, first by designing a fuzzy PID controller and considering the appropriate rules, we solved the problem of congestion control for the network and at the same time, we made optimal use of the existing network. The designed fuzzy PID controller showed how the desired queue length is well followed and by simulation, we showed that the output of the fuzzy PID controller is much better in terms of an output response, overshoot, speed and uncertainly than the classic controller. In the next step, which was the main topic of this research, it was suggested that the amount of packets loss is limited and is acceptable to an extent that does not exceed the threshold. If the packets loss increases too much, to overcome it, we designed a mechanism that includes different networks, each with its characteristics and designed with different capacities. If we use a network whose number of packets loss exceeds the allowable limit, with a management and supervisory system and using hybrid switching, it detects this issue and switches to a network with a higher capacity. By doing this, we always make sure that data subscribers never run out of bandwidth, and their satisfaction with the operators' increases.

7. REFERENCES

1. Clark, D., "The design philosophy of the DARPA internet protocols", ***ACM SIGCOMM Computer Communication Review***, Vol. 18, No. 4, (1988), 106–114. doi:10.1145/52325.52336
2. Floyd, S., and Jacobson, V., "Random early detection gateways for congestion avoidance", ***IEEE/ACM Transactions on Networking***, Vol. 1, No. 4, (1993), 397–413. doi:10.1109/90.251892
3. Christiansen, M., Jeffay, K., Ott, D., and Smith, F. D., "Tuning RED for Web traffic", ***IEEE/ACM Transactions on Networking***, Vol. 9, No. 3, (2001), 249–264. doi:10.1109/90.929849
4. Preye, U. H., and Nneka, O. L., "An Intelligent Fuzzy Logic System for Network Congestion Control", ***Circulation in Computer Science***, Vol. 2, No. 11, (2017), 23–30. doi:10.22632/ccs-2017-252-69
5. Domańska, J., and Domański, A., "Adaptive RED in AQM", In International Conference on Computer Networks, (2009), 174–183 Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-02671-3_21
6. Wu-chang Feng, Shin, K. G., Kandlur, D. D., and Saha, D., "The BLUE active queue management algorithms", ***IEEE/ACM Transactions on Networking***, Vol. 10, No. 4, (2002), 513–528. doi:10.1109/TNET.2002.801399
7. Hwang, L.-C., "M-GREEN: An active queue management mechanism for multi-QoS classes", ***Computer Standards & Interfaces***, Vol. 36, No. 1, (2013), 122–131. doi:10.1016/j.csi.2013.07.007
8. Feng, C.-W., Huang, L.-F., Xu, C., and Chang, Y.-C., "Congestion Control Scheme Performance Analysis Based on Nonlinear RED", ***IEEE Systems Journal***, Vol. 11, No. 4, (2017), 2247–2254. doi:10.1109/JSYST.2014.2375314
9. Bisoy, S. K., and Pattnaik, P. K., "An AQM Controller Based on Feed-Forward Neural Networks for Stable Internet", ***Arabian Journal for Science and Engineering***, Vol. 43, No. 8, (2018), 3993–4004. doi:10.1007/s13369-017-2767-9
10. Daigavhane, M. U., and Chawhan, M. D., "Congestion control algorithm for TCP in wireless network", In 2018 4th International Conference on Recent Advances in Information Technology (RAIT), (2018), 1–4. doi:10.1109/RAIT.2018.8388967
11. Alrshah, M. A., Al-Maqri, M. A., and Othman, M., "Elastic-TCP: Flexible Congestion Control Algorithm to Adapt for High-BDP Networks", ***IEEE Systems Journal***, Vol. 13, No. 2, (2019), 1336–1346. doi:10.1109/JSYST.2019.2896195
12. Tanenbaum, A. S., Computer Networks, (2003), Pearson Education.
13. Misra, V., Gong, W.-B., and Towsley, D., "Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED", In Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication - SIGCOMM '00, (2000), 151–160. doi:10.1145/347059.347421
14. Hollot, C. V., Misra, V., Towsley, D., and Wei-Bo Gong, "On designing improved controllers for AQM routers supporting TCP flows", Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat.

No.01CH37213), Vol. 3, (2001), 1726–1734. doi:10.1109/INFCOM.2001.916670

15. Barzamini, R., and Shafiee, M., "A new sliding mode controller for TCP congestion control", In 2011 19thTelecommunications Forum (TELFOR) Proceedings of Papers, (2011), 214–217. doi:10.1109/TELFOR.2011.6143529
16. Bisoy, S. K., and Pattnaik, P. K., "Design of feedback controller for TCP/AQM networks", ***Engineering Science and Technology, an International Journal***, Vol. 20, No. 1, (2017), 116–132. doi:10.1016/j.jestch.2016.10.002
17. Abualhaj, M. M., Abu-Shareha, A. A., and Al-Tahrawi, M. M., "FLRED: an efficient fuzzy logic based network congestion control method", ***Neural Computing and Applications***, Vol. 30, No. 3, (2018), 925–935. doi:10.1007/s00521-016-2730-9
18. Rezaee, A. A., and Pasandideh, F., "A Fuzzy Congestion Control Protocol Based on Active Queue Management in Wireless Sensor Networks with Medical Applications", ***Wireless Personal Communications***, Vol. 98, No. 1, (2018), 815–842. doi:10.1007/s11277-017-4896-6
19. Qu, S., Zhao, L., and Xiong, Z., "Cross-layer congestion control of wireless sensor networks based on fuzzy sliding mode control", ***Neural Computing and Applications***, Vol. 32, No. 17, (2020), 13505–13520. doi:10.1007/s00521-020-04758-1
20. Fan Yanfie, Ren Fengyuan, and Lin Chuang, "Design a PID controller for active queue management", In Proceedings of the Eighth IEEE Symposium on Computers and Communications. ISCC 2003, (2003), 985–990. doi:10.1109/ISCC.2003.1214244
21. Mohammadi, S., Pour, H. M., Jafari, M., and Javadi, A., "Fuzzy-based PID active queue manager for TCP/IP networks", In 10th International Conference on Information Science, Signal Processing and Their Applications (ISSPA 2010), (2010), 434–439, IEEE, 434–439. doi:10.1109/ISSPA.2010.5605462
22. Karam, Z. A., "Hybrid Fuzzy Congestion Controllers for Computer Networks Tuned by Modified Particle Swarm Optimization", ***International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems***, Vol. 7, No. 2, (2018), 17-26. doi:10.11601/ijates.v7i2.250
23. Taeib, A., Ltaeif, A., and Chaari, A., "A PSO Approach for Optimum Design of Multivariable PID Controller for nonlinear systems", In International Conference on Control, Engineering & Information Technology, Proceedings Engineering & Technology, (2013), 206–210.
24. Somwanshi, D., Bundele, M., Kumar, G., and Parashar, G., "Comparison of Fuzzy-PID and PID Controller for Speed Control of DC Motor using LabVIEW", ***Procedia Computer Science***, Vol. 152, (2019), 252–260. doi:10.1016/j.procs.2019.05.019
25. Zhen-Yu Zhao, Tomizuka, M., and Isaka, S., "Fuzzy gain scheduling of PID controllers", ***IEEE Transactions on Systems, Man, and Cybernetics***, Vol. 23, No. 5, (1993), 1392–1398. doi:10.1109/21.260670
26. Farshad Samadi, and Hamid Moghadam-Fard, "Active Suspension System Control Using Adaptive Neuro Fuzzy (ANFIS) Controller", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 28, No. 3, (2015), 396–401. doi:10.5829/idosi.ije.2015.28.03c.08
27. Kumar, P., and Kumar Chaudhary, S., "Stability and Robust Performance Analysis of Fractional Order Controller over Conventional Controller Design", ***International Journal of Engineering, Transaction B: Applications***, Vol. 31, No. 2, (2018), 322–330. doi:10.5829/ije.2018.31.02b.17
28. Kumar, V., and Patra, A., "Application of Ziegler-Nichols method for tuning of PID controller", ***International Journal of Electrical and Electronics Engineering***, Vol. 8, No. 2, (2016), 559–570

Persian Abstract

چکیده

یکی از مشکلات شبکه‌های امروزی سیستم انتقال آنهاست. در صورت پرشدن پهنای باند یک شبکه باید عوامل انسانی و فیزیکی دست بکار شده و جهت تامین پهنای باند آن ازیک سیستم انتقال جدید با ظرفیت بالاتر استفاده نمایند، که خیلی زمانبر و هزینه‌براست. در این مقاله روشی را پیشنهاد می‌دهیم تا علاوه بر استفاده بهینه از پهنای باند موجود در صورت پر شدن ظرفیت شبکه بصورت اتوماتیک به شبکه با پهنای باند بالاتر منتقل گردد. برای این منظور ابتدا با طراحی یک کنترل کننده PID فازی برای شبکه موجود سعی می‌شود کنترل ازدحام را انجام داده و حداکثر استفاده را از آن نمود. مشاهده می‌گردد که کنترل‌کننده پیشنهادی از نظر پاسخ خروجی، دنبال نمودن طول صف مطلوب، پایداری و عدم قطعیت در مقایسه با کنترل‌کننده کلاسیک بسیار بهتر عمل می‌نماید. درصورت افزایش داده‌های ورودی به شبکه تعداد بسته‌های بیشتری حذف می‌شوند و این باعث پایین آمدن کیفیت شبکه می‌شود. برای رفع این مشکل با استفاده از مدیریت پهنای باند، با در نظرگرفتن حد آستانه برای حذف بسته ها در هر شبکه درصورت عبور از این حد شبکه موجود به یک شبکه با ظرفیت بالاتر سوئیچ گردیده و مشکل پهنای باند و کیفیت شبکه حل می‌گردد و باعث رضایتمندی مشترکین می‌شود.

International Journal of Engineering

Journal Homepage: www.ije.ir

Designing Bi-directional Counters using Quantum-dot Cellular Automata Nanotechnology

Z. Amirzadeh[a], M. Gholami[b]

[a] *Department of Electrical Engineering, Mazandaran Institute of Technology, Babol, Iran*
[b] *Faculty of Technology and Engineering, University of Mazandaran, Babolsar, Iran*

PAPER INFO

Paper history:
Received 26 January 2020
Received in revised form 26 January 2021
Accepted 12 February 2021

Keywords:
Quantum-dot Celuular Automata
Counter
Bidirectional Counter

ABSTRACT

One of the major problems in designing highly compact integrated circuits is the power consumption of the circuits. Therefore, several technologies have been introduced to overcome the problems facing MOSFET technology. One of these technologies is the Quantum-Dot Cellular Atomata (QCA), which has several advantages. In this paper, we focus on computational logic gates based on the T-Latch circuit. T-latch is the basis of many circuit in arithmetic logic unit (ALU). The proposed structure for T-latch has a lower number of cells, occupied area and lower power consumption than existing methods. In the proposed T-Latch, compared to previous best designs, 6.45% cross section area and 44.49% power consumption were reduced. Also in this paper, for the first time a T-latch with reset terminal and a T-Latch with both set and reset terminals were designed. In addition, using the proposed T-latch, a 3-bit bidirectional up-down counter which consists of 204 quantum cells, 0.26 μm^2 cross-sectional area, delay of 5.25 clock cycles, a three-bit up-down counter with a reset pin and a three-bit up-down counter with set and reset terminals were made. The proposed up-down circuits are designed for the first time in QCA technology. All the design and simulation results are done in QCADesigner software.

doi: *10.5829/ije.2021.34.04a.16*

1. INTRODUCTION

In recent years, manufacturing of integrated circuits has improved with CMOS technology [1]. Reducing the size of transistors has reduced power consumption and increased the speed of integrated circuits [2], but due to the physical limitations of CMOS technology, ample research has begun to produce a new generation of integrated circuits. Among the emerging technologies, the technology of quantum cellular automata is more prominent than other technologies [3]. This technology has simple cells that are used as key elements in the creation of gates, wires and memories [4]. Also, while having a simpler structure, it has other capabilities, such as high speed and low power consumption, due to the small number of elements used in the circuit structure than complementary metal oxide semiconductor (MOS) technology [5]. It also looks like this technology will be a new way to implement digital circuits in the future [6].

Counters are the most common circuits in the processing unit, and these circuits are generally sequential logic circuits that are activated by an external signal called a clock signal [7]. In the previous works, counters have been introduced in QCA technology, each of which has its advantages and disadvantages. Amirzadeh and Gholami [8] proposed a 3-bit D-type flip-flop based counter that has relatively large cell counts and cross-sections area, and high latencies despite proper performance. It has also been suggested Majeed et al. [9] T-type flip-flop based counters that, despite improvements in area level and delay rate, they have a large number of cells. In addition, none of the counters proposed in the past were able to count specific numbers (for example for counting between 2 and 5) due to lack of set and reset terminals. A reset-based counter using D-type flip-flop was proposed by Zoka and Gholami [10], which had a relatively large number of cells, cross-sectional area, delay, and power consumption. Therefore, in this paper, using the proposed T-latch and T-type flip-

*Corresponding Author Email: m.gholami@umz.ac.ir (M. Gholami)

Please cite this article as: Z. Amirzadeh, M. Gholami, Designing Bi-directional Counters using Quantum-dot Cellular Automata Nanotechnology, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 901-907

flop structure, bidirectional counters, bidirectional counters with reset capability and bidirectional counters with set and reset capabilities will be presented.

This article will provide the following sections. In the second part, the basic concepts and how to make quantum cells in QCA technology will be discussed. In the third section, the proposed conventional up-down counter circuit and the proposed counter with set and reset terminals in QCA technology are reported. Also, simulation results which shows the performance of the proposed circuits are presented in this section. Finally, power analysis are discussed in section four.

2. BASICS OF QCA

Quantum-dot cellular automata (QCA) is an emerging technology for the design of nanoscale electronic devices that meet the requirements of modern digital systems [11]. In this technology, states are not defined by voltage levels, but by the position of the electrons [12]. In QCA, a cell has four quantum dots located at the corners of a square [13]. Within each cell, there are two free electrons. Because of the electron repulsion between them, these electrons are separated at the farthest distance [14]. Using these cells, different types of QCA based circuits can be made. One of the basic logic gates in the QCA is the majority gate. By using the majority gate and fixing one of the inputs, the AND and OR logic function can be created [15]. Another basic logic gate in a quantum cellular atom is the NOT gate that is obtained by placing the cells obliquely above or below the target cell [16]. Another important block in the QCA is the wire, which is created by putting cells together in a linear fashion. The QCA also uses a clock system to control information from the QCA, which includes four phases Switch, Hold, Release, and Relax.

In the switch phase, the electron wants to get to the new state from the previous state. In the hold phase, the electrons are not affected by the electrons of the adjacent cells and the cell retains its polarity. In the release phase, the energy of the inner points of the cell decreases to a point that the cell loses its polarity. In the relax phase, there are no internal points and the cell does not have any effect on its adjacent cells.

Also, the interdot potential barriers of a QCA cell are induced by electric or magnetic field. So, a common approach to realizing clock circuits is to bury the wires below the QCA level [17].

3. PROPOSED UP-DOWN COUNTERS IN QCA NANOTECHNOLOGY AND SIMULATION RESULTS

Bidirectional counters have the ability to count up and down directions for any given sequence. It is also possible to reverse the count at any point in the counting sequence, using an additional control input [18]. In the following, the proposed bidirectional counters will be discussed and their principles and work will be examined.

The main blocks of the proposed bidirectional counter circuits are T flip-flops and 2×1 multiplexer. In this paper, for the design of the proposed bidirectional counters, proposed T flip-flop and proposed T flip-flop with set and reset terminals are used which are appropriate in terms of cell number of cells, occupied area, and design capability. The proposed bidirectional counter circuits will be designed with minimum cell number and delay while also have reset or set ability. In the proposed designs 2×1 multiplexer is used based on the circuit in literature [19].

One of the most noteworthy topics in circuit design is the memory elements and flip flops that have been considered in QCA technology. T flip-flop is one of these flip flops that is used for toggling an input. T flip-flop is works as follows: if T='1' and in the edge of the clock, the output will reverse the previous state and in all other cases, the output will be the same as its previous state. This is true for flip-flops sensitive to edge of the clock (rising or falling).

A proposed bi-directional counter which counts ascending or descending numbers in QCA nanotechnology is shown in Figure 1. This counter is asynchronous thee-bit up-down counter and is designed using proposed T-flip flops. This counter will counts between 0 (000) and 7 (111). When the down counter's selector is activated, the counter start counting reversely (for example from 7 to 0). So the counting string in this counter, in the up counting mode is 0, 1, 2, 3, 4, 5, 6, 7 and in the down counting mode, it is 7, 6, 5, 4, 3, 2, 1, 0. It should be mentioned that in this block diagram, when both of the selectors are zero, the counter will be UP-counter and when both of the selectors are logical one, the counter will be down-counter. In addition, it should be noted that when the selector's values are being changed, the counting process will start from the same point in reverse manner.

In this design, to have the correct counting numbers for proposed bidirectional counter and also, in order to be able to reverse the count at any point, the circuit diagram block was designed in a way that: Since the circuit is asynchronous and the flip-flops are sensitive

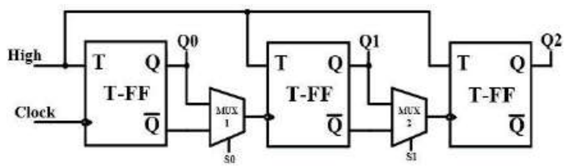


Figure1. Logical diagram of proposed three-bit bidirectional counter

to the downward edges, for counting incremental numbers, the Q output of each flip flop must go to the clock input of next flip-flop, and when the circuit wants to count down, the $\bar{Q}$ output of the each flip flop should be connected to the clock input next flip-flop. Therefore, a multiplexer is used at the output of each flip-flops as shown in Figure 1. As can be seen, the inputs of multiplexer come from outputs of Q and $\bar{Q}$ of flip-flops in previous stages. When the selector input of multiplexer is zero, the Q will be sent to the clock of next stage and the circuit will be up-counter and when the selector is zero, the circuit acts as down-counter since the $\bar{Q}$ will be sent to the clock of next stage. With this logical block diagram, this circuit starts counting from zero to 7 and then starts again to 7 counting and counts to 0 at any point in time that the selectors change from zero to logical one. This is obviously illustrated in Figure 1.

Figure 2 shows the proposed three-bit bidirectional counter structure in QCA technology, which consists of 204 quantum cells, 0.26 μm^2 cross-sectional area, delay of 5.25 clock cycles. The proposed structure consists of three proposed T-latches, three level to edge converters and two 2 × 1 multiplexers which is designed according to structure of Figure 1.

Figure 3 shows the simulation results of the proposed bidirectional counter in QCA technology and works as follows: When the first and second selectors are zero, the up-counting is activated and counter counts from 0 to 7 and every time the first and second selectors

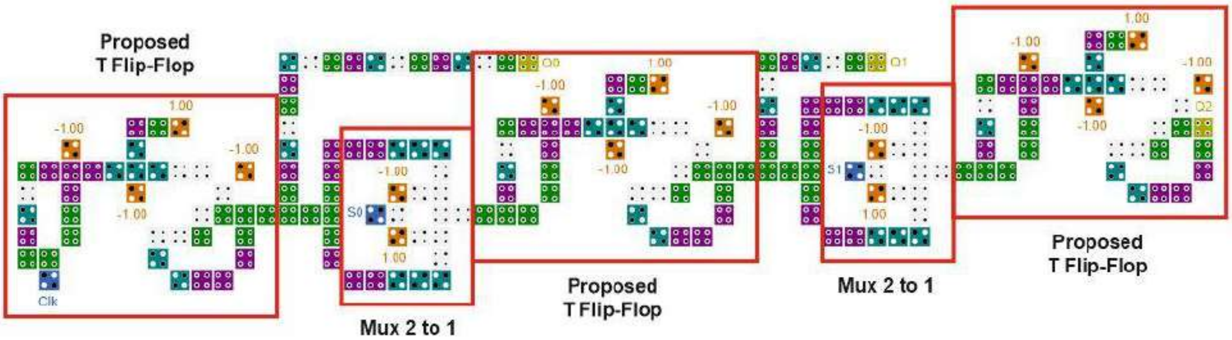


Figure 2. Layout of proposed bidirectional counter

are logical one, the down-counting mode will be activated and the proposed counter starts counting from 7 to 0. As a result, when the first and second selectors are zero and according to the initial conditions given to the circuit, the circuit begins to count upwards and when the selectors become logical one, the counter starts counting backwards at any point in the count. Figure 3 shows the correct operation of the circuit.

Figure 4 shows the block diagram of the proposed bidirectional counter with a reset pin. In this design, to make the circuit correctly counts, and after counting it is possible to reverse the count at any point in the counting sequence and perform the reset operation correctly in the circuit, the circuit block diagram is thus designed according to Figure 4. Based on to this figure the proposed circuit work as follows. When the circuit is

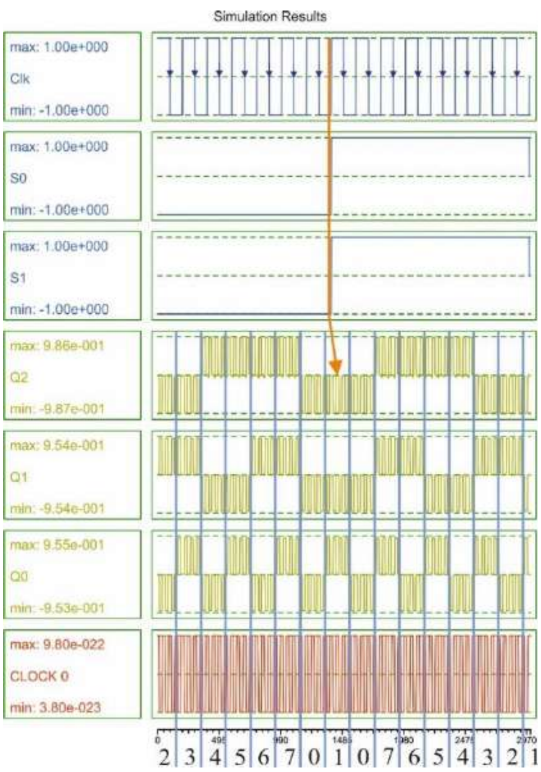


Figure 3. Output waveforms of proposed bidirectional counter

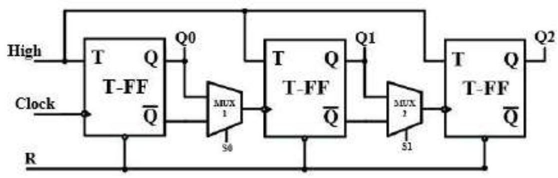


Figure 4. Logical block diagram of proposed bidirectional counter with reset ability

asynchronous and the clock is sensitive to the downward edges, the Q (output of each stages) must go to the clock of the next stage in order to count the upward numbers, and when the circuit wants to count downward, the $\bar{Q}$ must be sent to the clock of the next stage. Therefore, again multiplexer is inserted between each flip-flops to do these acts. As a result, when the first and second selectors are zero and according to the initial conditions given to the circuit, the circuit begins to count upwards and when the selectors become logical one, the counter starts counting backwards at any point in the count. In addition, since this circuit has a reset input, when the reset is activated the output will be zero in any time of counting.

Figure 5 shows the proposed three-bit bidirectional counter with reset ability in QCA technology, comprising 288 cells, 0.36 μm^2 cross-sectional area and 5.75 clock cycles delay. The proposed structure consists of three proposed T-latches with reset input, three level to falling edge converter and two 2 × 1 multiplexers according to topology of Figure 4.

Figure 6 shows the simulation results of the proposed bidirectional reset-based counter in QCA technology and works as follows: When the first and

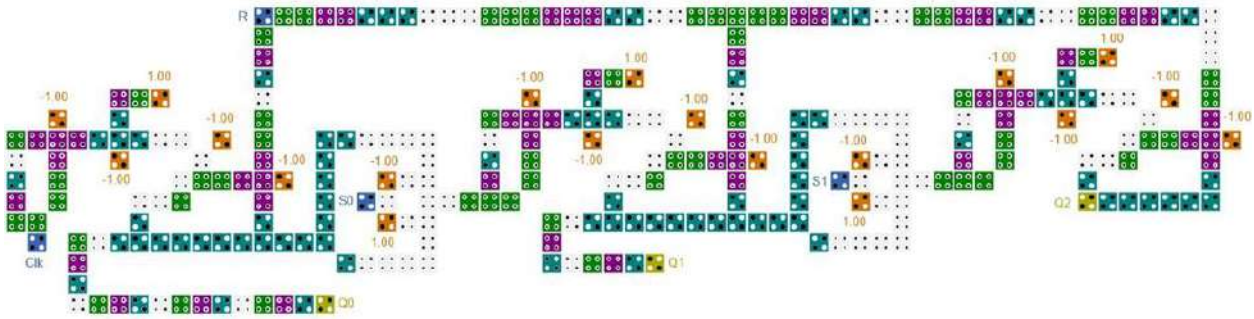


Figure 5. Layout of proposed bidirectional counter with reset ability

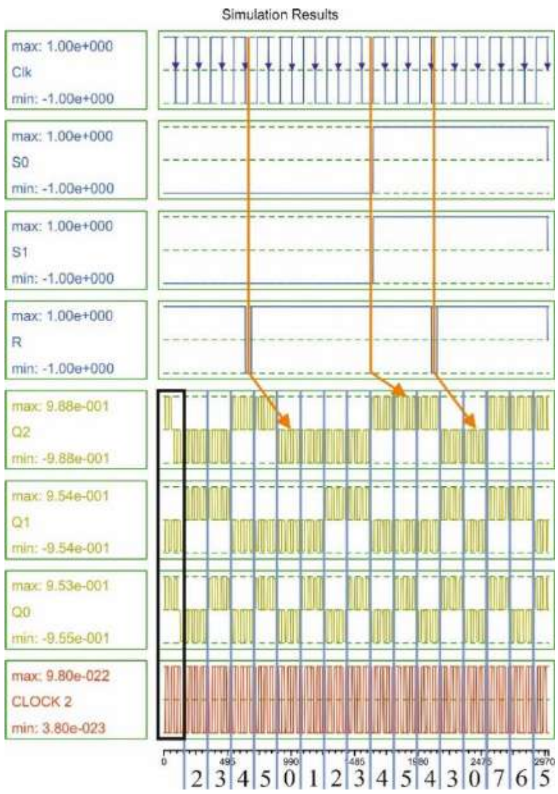


Figure 6. Output waveforms of proposed bidirectional counter with reset ability

second selectors are logical zero, the up-counting is activated and started and when the first and second selectors are logical one, the down-counting will be activated. In addition, since this circuit has a reset input, when the reset is activated the output will be zero in any time of counting. All of these can be seen in Figure 6. Figure 7 shows the block diagram of the proposed bidirectional counter with set and reset inputs. The performance of this circuit is similar to the previous designs and only set and reset capabilities have been added to the proposed counter. With this structure we can have an up-down counter with the ability to count arbitrarily. The counting can be continued in these counters while counting using set and reset inputs from the number that we want.

Figure 8 illustrates the proposed three-bit bi-directional counter-set structure with set and reset inputs in QCA technology, comprising 363 cells, 0.40 μm^2 cross section area, and 5.75 clock cycles delay. The proposed structure consists of three proposed T-latches with a set and reset pins, three level to edge converters and two 2×1 multiplexers designed according to Figure 7. In addition, Figure 9 illustrates the simulation results of the three-bit bi-directional set-counter and the proposed reset in QCA technology. In this figure the process of counting is separated for two cases down and up counting.

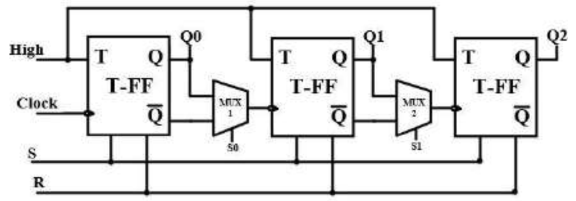


Figure 7. Logical diagram of proposed bidirectional counter with set and reset abilities

Table 1 shows the comparison of proposed designs. Since, bidirectional counter is designed for first time in this paper, we cannot find any related works in this domain for comparing the results.

4. POWER SIMULATIONS

Recent studies show that although there is no electricity in QCA gates, these devices are not without energy loss [20]. The results show that the number of circuit inputs and the geometric compression in QCA are two very effective factors in energy loss [21]. All the proposed designs have been investigated at three tunneling levels $0.5E_K$, $1E_K$ and $1.5E_K$ at T=2ºK. The results are abbreviated in Table 2. Also, all the relationships related to how to calculate energies are discussed by Toress et al. [22]. Figures 10 and 11 show the power consumption of the proposed bidirectional counter structure, the proposed bidirectional counter with reset ability and the

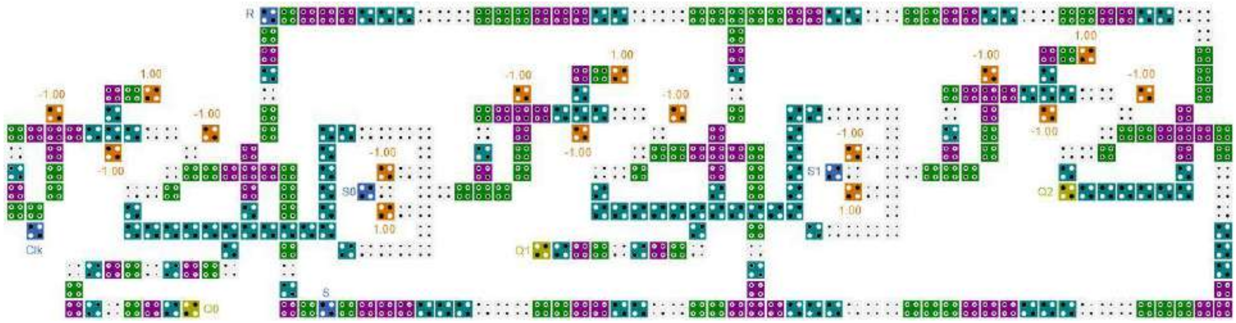


Figure 8. Layout of proposed bidirectional counter with set and reset abilities

(a) (b)

Figure 9. Output waveforms of proposed bidirectional counter with set and reset abilities (a) up direction; (b) down direction

TABLE 1. Results of proposed designs

Circuit	Cell count (# Cells)	Area (μm^2)	Latency ($10^{-12}s$)	Total Energy dissipation 0.5 E_k (meV)	Set input	Reset input
Proposed T-Latch	21	0.0174	0.75	16.22	NO	NO
Proposed T-Latch with Reset	31	0.03	1.25	18.86	NO	YES
Proposed T-Latch with Set and Reset	36	0.04	1.25	22.53	YES	YES
Proposed up-down counter(Fig.2)	204	0.26	5.25	117.77	NO	NO
Proposed up-down with reset (Fig.5)	288	0.36	5.75	213.16	NO	YES
Proposed up down with set and reset (Fig.8)	363	0.40	5.75	167.21	YES	YES

TABLE 2. Results of average leakage energy dissipation and average switching energy dissipation

Circuit	Average Leakage energy dissipation (meV)			Average Switching energy dissipation (meV)		
	0.5 E_K	**1.0 E_K**	**1.5 E_K**	**0.5 E_K**	**1.0 E_K**	**1.5 E_K**
Proposed T-Latch	8.36	22.16	36.89	7.86	6.60	5.42
Proposed T-Latch with Reset	11.79	31.69	53.27	7.07	5.93	4.91
Proposed T-Latch with Set and Reset	13.41	36.02	60.64	9.12	7.55	6.22
Proposed up-down counter(Fig.2)	84.68	224.54	377.53	33.09	27.32	22.61
Proposed up-down with reset (Fig.5)	102.62	293.05	508.96	110.54	90.55	73.96
Proposed up down with set and reset (Fig.8)	127.13	366.68	639.28	40.08	33.28	27.54

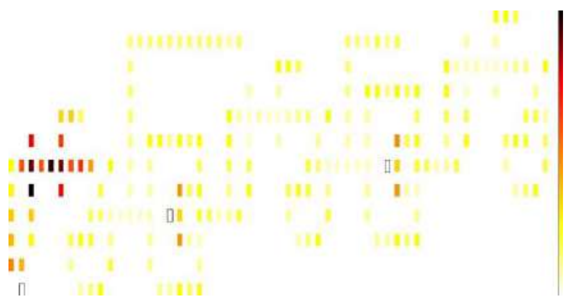


Figure 10. The power dissipation maps of proposed up-down counter

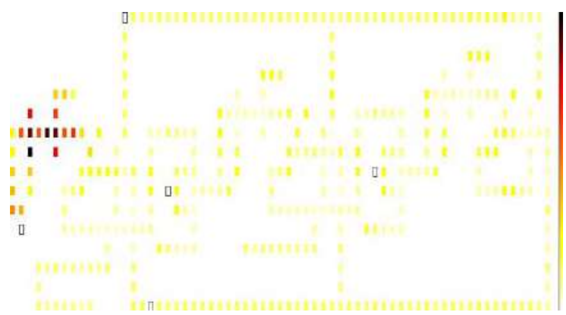


Figure 11. The power dissipation maps of proposed up down with set and reset

proposed up-down counter with set and reset abilites at the 0.5E_K level, respectively. In addition, the results of average leakage energy dissipation and average switching energy dissipation of proposed designs are abbreviated in Table 2.

5. CONCLUSION

In this paper, we discuss the basic concepts of quantum cells and schemes of proposed T flip-flop, T flip-flop with reset and T flip-flop with set and reset. In addition three Up-Down counters are proposed: conventional bidirectional counter, bidirectional counter with reset ability and bidirectional counter with set and reset inputs. Also, the proposed design are simulated using QCADesigner and QCAPro.

6. REFERENCES

1. Purkayastha, T., De, D. and Chattopadhyay, T., "Universal shift register implementation using quantum dot cellular automata", ***Ain Shams Engineering Journal***, Vol. 9, No. 2, (2018), 291-310. https://doi.org/10.1016/j.asej.2016.01.011
2. Alamdar, H., Ardeshir, G. and Gholami, M., "Novel quantum-dot cellular automata implementation of flip-flop and phase-frequency detector based on nand-nor-inverter gates", ***International Journal of Circuit Theory and Applications***, Vol. 49, No. 1, (2021), 196-212. https://doi.org/10.1002/cta.2825

3. Roshan, M.G. and Gholami, M., "Novel d latches and d flip-flops with set and reset ability in qca nanotechnology using minimum cells and area", ***International Journal of Theoretical Physics***, Vol. 57, No. 10, (2018), 3223-3241. https://doi.org/10.1007/s10773-018-3840-1
4. Binaei, R. and Gholami, M., "Design of multiplexer-based d flip-flop with set and reset ability in quantum dot cellular automata nanotechnology", ***International Journal of Theoretical Physics***, Vol. 58, No. 3, (2019), 687-699. https://doi.org/10.1007/s10773-018-3967-0
5. Gholami, M., Binaei, R. and Gholamnia Roshan, M., "Novel phase-frequency detector based on quantum-dot cellular automata nanotechnology", ***International Journal of Engineering***, Vol. 33, No. 2, (2020), 269-276. https://doi.org/10.5829/IJE.2020.33.02B.11
6. Niknezhad Divshali, M. and Rezai, A., "Novel circuits design for siso shift register in qca technology", ***Journal of Circuits, Systems and Computers***, (2021). https://doi.org/10.1142/S0218126621502030
7. Moharrami, E. and Navimipour, N.J., "Designing nanoscale counter using reversible gate based on quantum-dot cellular automata", ***International Journal of Theoretical Physics***, Vol. 57, No. 4, (2018), 1060-1081. https://doi.org/10.1007/s10773-017-3638-6
8. Amirzadeh, Z. and Gholami, M., "Counters designs with minimum number of cells and area in the quantum-dot cellular automata technology", ***International Journal of Theoretical Physics***, Vol. 58, No. 6, (2019), 1758-1775. https://doi.org/10.1007/s10773-019-04070-2
9. Majeed, A.H., Alkaldy, E., bin Zainal, M.S. and Nor, B.M., "Synchronous counter design using novel level sensitive t-ff in qca technology", ***Journal of Low Power Electronics and Applications***, Vol. 9, No. 3, (2019), 27. https://doi.org/10.3390/jlpea9030027
10. Zoka, S. and Gholami, M., "A novel rising edge triggered resettable d flip-flop using five input majority gate", ***Microprocessors and Microsystems***, Vol. 61, (2018), 327-335. https://doi.org/10.1016/j.micpro.2018.06.006
11. Abutaleb, M., "A novel configurable flip flop design using inherent capabilities of quantum-dot cellular automata", ***Microprocessors and Microsystems***, Vol. 56, (2018), 101-112. https://doi.org/10.1016/j.micpro.2017.11.003
12. Binaei, R. and Gholami, M., "Design of novel d flip-flops with set and reset abilities in quantum-dot cellular automata nanotechnology", ***Computers & Electrical Engineering***, Vol. 74, (2019), 259-272. https://doi.org/10.1016/j.compeleceng.2019.02.002
13. Alamdar, H., Ardeshir, G. and Gholami, M., "Phase-frequency detector in qca nanotechnology using novel flip-flop with reset terminal", ***International Nano Letters***, Vol. 10, (2020), 111-118. https://doi.org/10.1007/s40089-020-00300-2
14. Jeon, J.-C., "Low-complexity qca universal shift register design using multiplexer and d flip-flop based on electronic correlations", ***The Journal of Supercomputing***, Vol. 76, No. 8, (2020), 6438-6452. https://doi.org/10.1007/s11227-019-02962-y
15. Roshan, M.G. and Gholami, M., "4-bit serial shift register with reset ability and 4-bit lfsr in qca technology using minimum number of cells and delay", ***Computers & Electrical Engineering***, Vol. 78, (2019), 449-462. https://doi.org/10.1016/j.compeleceng.2019.08.002
16. Amirzadeh, Z. and Gholami, M., "Analysis and design of the pseudo-random bit generator in the technology of quantum-dot cellular automata", ***International Journal of Theoretical Physics***, Vol. 59, No. 1, (2020), 29-48. https://doi.org/10.1007/s10773-019-04262-w
17. Campos, C.A.T., Marciano, A.L., Neto, O.P.V. and Torres, F.S., "Use: A universal, scalable, and efficient clocking scheme for qca", ***IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems***, Vol. 35, No. 3, (2015), 513-517. https://doi.org/10.1109/TCAD.2015.2471996
18. Sasamal, T., Gaur, H., Singh, A. and Mohan, A., Novel approaches for designing reversible counters, in Design and testing of reversible logic. 2020, Springer.37-48. https://doi.org/10.1007/978-981-13-8821-7_3
19. Asfestani, M.N. and Heikalabad, S.R., "A unique structure for the multiplexer in quantum-dot cellular automata to create a revolution in design of nanostructures", ***Physica B: Condensed Matter***, Vol. 512, (2017), 91-99. https://doi.org/10.1016/j.physb.2017.02.028
20. Angizi, S., Moaiyeri, M.H., Farrokhi, S., Navi, K. and Bagherzadeh, N., "Designing quantum-dot cellular automata counters with energy consumption analysis", ***Microprocessors and Microsystems***, Vol. 39, No. 7, (2015), 512-520. https://doi.org/10.1016/j.micpro.2015.07.011
21. Rahimi, E., "Energy dissipation of quantum-dot cellular automata logic gates", ***Micro & Nano Letters***, Vol. 11, No. 7, (2016), 369-371. https://doi.org/10.1049/mnl.2015.0535
22. Torres, F.S., Wille, R., Niemann, P. and Drechsler, R., "An energy-aware model for the logic synthesis of quantum-dot cellular automata", ***IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems***, Vol. 37, No. 12, (2018), 3031-3041. https://doi.org/10.1109/TCAD.2018.2789782

Persian Abstract

چکیده

یکی از مشکلات مهم در طراحی مدارهای مجتمع بسیار فشرده، کاهش توان مصرفی مدارها است. بنابراین تکنولوژی های متعددی برای خروج از مشکلات پیش روی تکنولوژی MOSFET معرفی شده است. یکی از این تکنولوژی‌ها آتوماتای سلولی کوانتومی نقطه ایی (QCA) است، که دارای مزایای متعددی میباشد. در این مقاله بروی گیت های محاسباتی مدار منطقی، لچ T تمرکز گردید که به عنوان مدار پایه در بسیاری از مدارات دیگر و همچنین واحدهای محاسباتی و مقایسه ای استفاده می گردد. ساختار پیشنهاد شده از سلول، مساحت اشغال شده و توان مصرفی کمتری نسبت به روشهای پیاده‌سازی موجود برخوردار است. بطوریکه در لچ T پیشنهادی در مقایسه با بهترین طرحهای گذشته، ٪6.45 سطح مقطع و ٪44.49 توان مصرفی کاهش داده شد. همچنین در این مقاله، برای اولین بار لچ T با پایه ریست و لچ T با پایه ست و ریست نیز طراحی شد. علاوه بر این‌ها در ادامه با استفاده از لچ T پیشنهادی، شمارنده سه بیتی دوجهته (بالا و پایین شمار) که شامل 204 سلول، $0.26\ \mu m^2$ سطح مقطع، 5.25 سایکل کلاک تاخیر، شمارنده سه بیتی دوجهته با پایه‌ی ریست و شمارنده سه بیتی دوجهته با پایه‌ی ست و ریست برای اولین بار ساخته شد که با بررسی هایی که صورت گرفته مقالات مشابهی که اعداد را به صورت دوجهته بشمارند وجود نداشته است. تمامی طراحی‌ها و نتایج شبیه سازی‌ها در نرم افزار QCADesigner انجام شده است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Cogging Torque Minimization in Transverse Flux Permanent Magnet Generators using Two-step Axial Permanent Magnet Segmentation for Direct Drive Wind Turbine Application

R. Nasiri-Zarandi*[a], A. Mohammadi Ajamloo[b], K. Abbaszadeh[b]

[a] *Electrical Machine Group, Niroo Research Institute (NRI), Tehran, Iran*
[b] *Department of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran*

PAPER INFO

Paper history:
Received 19 June 2019
Received in revised form 18 October 2020
Accepted 04 November 2020

Keywords:
Cogging Torque
Design of Experiment
Magnetic Equivalent Circuit
Permanent Magnet
Permanent Magnet Segmentation
Taguchi Method
Transverse Flux Machine
Wind Turbine

ABSTRACT

Transverse flux permanent magnet machines (TFPMs) are categorized as synchronous machines that benefit from having high value of torque density and capablity of accommodating high pole numbers. These characteristics make TFPMs suitable candidates for low-speed applications where high torque density value is requred such as direct drive wind turbine application. Despite the aforementioned advantages, TFPMs suffer from intrinsically high cogging torque value which is an important concern for wind turbine application. This paper focuses on axial PM segmentation technique to minimized cogging torque of TFPM topologies. Concept of the proposed method is discussed using analytical equations and optimum segmentaion angle is formulized. Non-linear magnetic equivalent circuit (MEC) is adopted where the PM segmentation, armature reaction, rotor transition and iron saturation effect are carefully modeled. The results of the MEC simulation are compared with the finite element method (FEM) results in terms of accuracy and computational time. The results from the analysis indicate that the proposed MEC method is almost ten times faster than FEM with reasonable level of precision. Taguchi method is adopted as a fast-response optimization method to improve the generator torque characteristics. The results show that the cogging torque has reduced by 97% with respect to the initial design while the average torque has only dropped by 8% which is an acceptable side effect due to the significant improvement in machine cogging torque.

doi: *10.5829/ije.2021.34.04a.17*

1. INTRODUCTION

The gearbox and transmission equipment in wind turbine systems reduce the efficiency of the wind generation system [1]. Furthermore, the gearing system reduces the reliability of the wind turbine and requires periodic maintenance [2, 3]. The absence of transmission equipment is considered an advantage for the direct-drive wind turbines (DDWTs), however, some considerations need to be taken into account. Owing to the elimination of the gearing system in DDTWs, the generator rotational speed is reduced with respect to the geared systems. Generally, in rotating electrical machines, the speed and volume are inversely proportional (at constant electrical and magnetic loadings). Consequently, in DDWT systems, the volume and weight of the generator increase, significantly. These problems result in increased generator manufacturing cost and stronger and expensive tower structure. Therefore, a generator with a high value of torque density is needed in DDWT systems [4]. In wind turbine systems, especially in DDWTs, generator cogging torque is an important criterion that must be taken into consideration. This is due to the fact that high values of the cogging torque prevent the turbine from starting up in low-speed wind [5]. Additionally, cogging torque and torque ripple induce noticeable noise and vibration to the structure [6–8]. So, based on the aforementioned discussion, two crucial characteristics of DDWT generators are: having high-value torque density and low value of cogging torque.

*Corresponding Author Institutional Email: *RNasiriZarandi@nri.ac.ir* (R. Nasiri-Zarandi)

Please cite this article as: R. Nasiri-Zarandi, A. Mohammadi Ajamloo, K. Abbaszadeh, Cogging Torque Minimization in Transverse Flux Permanent Magnet Generators using Two-step Axial Permanent Magnet Segmentation for Direct Drive Wind Turbine Application, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 908-918

Transverse flux permanent magnet machines (TFPMs) are categorized as synchronous machines which benefit from having a high value of torque density and capability of accommodating a high number of poles. These characteristics make TFPMs potential candidates for DDWT systems [9]. Generally, in TFPMs, the number of rotor poles are integer multiples of stator teeth which results in a considerably high value of cogging torque [10]. To this end, cogging torque minimization of TFPMs is discussed in many kinds of researches. Skewing stator U-I cores are investigated in [11] and effective results are presented. However, considering the laminated structure of U-I cores, shaping the proposed skewed parts is a very difficult procedure in practice. Optimizing the PM dimension is carried out in [12] to reduce the cogging torque of the tubular TFPM. However, the cogging torque reduction resulted from this technique is very weak. Unequal pole pitch of the rotor poles is studied in [13] and a noticeable reduction in cogging torque along with improvement in voltage harmonic components is achieved. However, the average electromagnetic torque has been reduced in this technique. Using different Halbach–Array structures to minimize the cogging torque is studied in [14, 15] and promising results are presented. However, the manufacturing process of Halbach-Array TFPM is relatively complicated. Shifting and unequal width of stator teeth is discussed in [16] which effective results along with experimental validations are presented. However, the proposed technique is not applicable to all TFPM topologies and is mainly applicable to certain topologies which use expensive SMC cores.

Generally, three phases TFPMs are built in three separate modules where each module belongs to a certain phase. In this case, three separate phases with almost zero magnetic couplings among them are developed. The stator of the TFPM topology is mainly developed based on using U and I-shaped laminated cores. The rotor of the earliest generation of the TFPMs contained two rows of PMs in each phase placed on an iron yoke. The stator was composed of only U-shaped laminated cores surrounding a circumferential winding [17]. The problem was that only half of the PMs contributed to useful power production while the other half created leakage flux and even reduced the useful power [17]. The above-mentioned deficiency was partially overcome by adding I-shaped laminated cores beside U-cores. By adding the I-shaped cores, the other half of PMs were shunted and didn’t create the leakage flux, but still, these PMs couldn’t produce useful power. To bring all the PMs into process of useful power production, the rotor yoke was separated into upper and lower rows [18]. By applying this modification, full utilization of PMs was achieved since all PMs were brought into the main flux path as shown in Figure 1. Regarding this figure which presents one pole pair flux path of the TFPM, it is seen that the

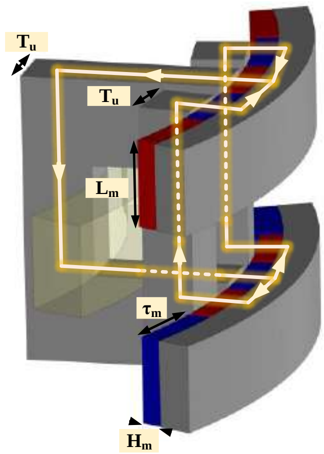


Figure 1. Main flux path of a single-phase TFPM in one pole pair

flux path has three dimensional (3D) nature. Therefore, analysis of TFPM topologies requires the 3D finite element method (FEM). FEM is a popular solution with a high level of accuracy and is extensively used in 2D and 3D analysis [19, 20]. However, applying 3D FEM is a time taking process especially for optimization purposes. Thus, many papers investigated using alternative analysis tools to substitute for the time-consuming FEM. The magnetic equivalent circuit (MEC) method has been considered as a potential substitute for the FEM due to its fast computational process and acceptable level of accuracy [21, 22]. It needs to be stated that getting an acceptable level of accuracy from the MEC method requires precise modeling of iron non-linearity, fringing flux, air gap permeances, and rotor transition which are taken into account in this study.

In this paper, aiming to minimize the cogging torque of a 2 kW TFPM, a two-step PM segmentation technique is proposed. Axial PM segmentation concept is explained and the related formulation is extracted to achieve a global formula for finding the optimum segmentation angle. To reduce the processing time of the analysis, a non-linear dynamic MEC model is developed and its results are validated by FEM. To improve the torque characteristics (electromagnetic and cogging torques) of the TFPM, fast-computing design of experiments (DOE) optimization is used based on the Taguchi method.

2. PM SEGMENTATION CONCEPT IN TFPMS

Ideally, three phase TFPMs are considered as three separate modules with no magnetic coupling among the phases (modules). However, small magnetic leakage flux between modules may exist resulting a magnetic coupling between the adjacent modules. However, this leakage flux is very weak which can be neglected in the analysis. So, by assuming three magnetically decoupled modules, cogging torque of Three-phase TFPMs is calculated for each module (i.e. phase A), independently. Then, cogging torque waveforms of phase B and C are obtained by shifting the cogging torque waveform of phase A by ±120° electrical degree. Consequently, the cogging torque of a three-phase TFPM is developed as in Equation (1);

$$T_{cog} = \sum_{n=1}^{\infty} a_n[\sin(n\theta_e) + \sin(n(\theta_e + \frac{2\pi}{3})) + \sin(n(\theta_e - \frac{2\pi}{3}))] \quad (1)$$

Here, the cogging torque of each phase is written in the form of the Fourier series where the first, second, and third terms belong to phase A, B, and C respectively.

So, it can be understood from Equation (1) that in a TFPM with three separate phases, all the cogging torque harmonic components are canceled in a three-phase cogging torque waveform except multiples of the third component. Actually, multiples of the third components due to obtaining the same phase are tripled in the three-phase waveform. Evidently, in a three-phase cogging torque waveform, the major harmonic component is one of the multiples of the third harmonic component. So, identifying the major harmonic component and minimizing it is pursued. To achieve such a goal, the technique of axial PM segmentation is proposed.

In this regard, Figure 2 represents the PM arrangement of TFPM before and after PM segmentation. According to Figure 2(b), two-step PM Segmentation is adopted where each PM is split into two upper and lower PMs which are shifted by a specific angle (φ_m) with respect to upper PM. Consequently, a single-phase cogging torque waveform could be written in a Fourier series of Equation (2).

$$T_A = \sum_{n=1}^{\infty}[a_n \sin(n\theta_e) + a_n' \sin(n(\theta_e + \frac{p}{2}\varphi_m))] \quad (2)$$

Here, the first term represents the effect of the upper PMs and the second term represents the lower PMs which are shifted by an angle of φ_m with respect to upper PMs. Here, p is the number of poles.

Regarding Equation (2), if the effects of fringing and leakage fluxes are neglected, an and a_n' can be assumed equal. So, it is possible to select a desired harmonic component from the single-phase cogging torque waveform and eliminate it by adjusting the shifting angle (φ_m). It means that the dominant component of the three-phase cogging torque waveform (i.e. ith component) could be theoretically suppressed by choosing a specific shift angle as in Equation (3);

$$T_i = a_i \sin(i\theta_e) + a_i \sin(i(\theta_e + \frac{p}{2}\varphi_m)) = 0 \quad \Rightarrow \varphi_m = \frac{360^0}{ip} \quad (3)$$

To analyze the proposed solution, a 2 kW TFPM generator with a rated speed of 350 rpm is designed for the DDWT application where its specifications are listed in Table 1. The detailed design procedure is presented in the next section. In Figure 3, cogging torque waveforms of single-phase and three-phase TFPM are depicted. As stated earlier, the three-phase cogging torque is developed by adding three single-phase waveforms that have 120° electrical degree shift angles.

Harmonic components of the single-phase and three-phase cogging torque waveforms are extracted and shown in Figure 4. Accordingly, the 2nd, 4th, and 6th components are the dominant components in single-phase cogging torque waveform. The 8th, 10th, and 12th components are minor harmonic components while the other harmonics components are negligible.

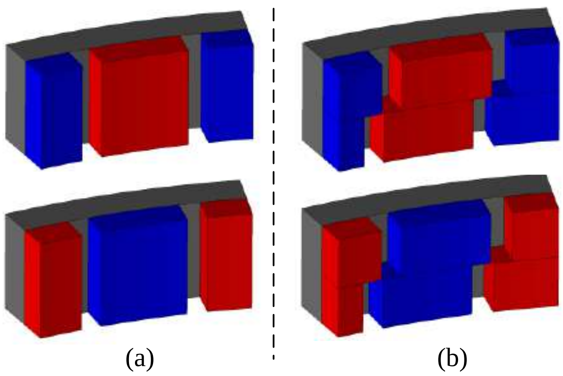


Figure 2. One pole pair view of PM configurations in (a) conventional TFPM topology, and (b) the proposed two-step axially segmented TFPM topology

TABLE 1. Design specifications of proposed TFPM

Parameter	Symbol	Value	Unit
Output power	P	2000	W
Nominal speed	n	350	rpm
No. of phases	m	3	---
No. of poles	p	40	---
Air gap diameter	D_g	203	mm
Rotor outer diameter	D_o	225	mm
Machine length	L	180	mm
Air gap length	g	1	mm
Stator shoe width	T_u	10.5	mm
PM length	L_m	15	mm
PM thickness	d_m	4	mm
Number of armature turns per phase	N	158	---
PM coercive force	H_c	-880	kA/m
PM residual flux density	B_r	1.25	T

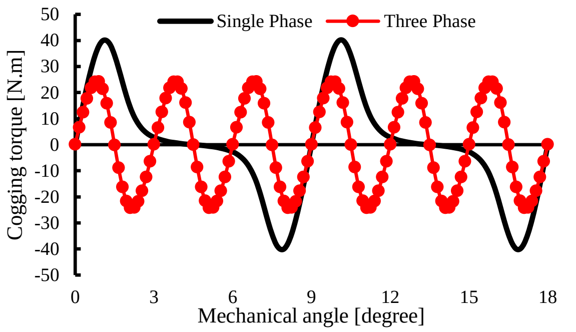


Figure 3. Single-phase and three-phase cogging torque waveforms of the TFPM generator

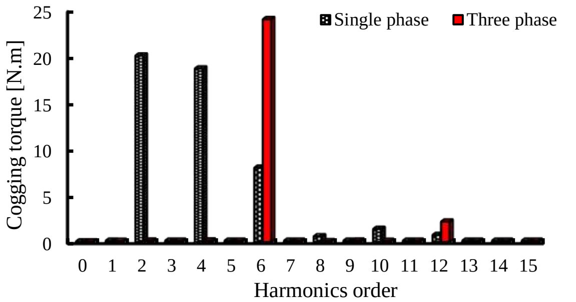


Figure 4. Cogging torque harmonics components of single-phase and three-phase TFPM generator

Since all the harmonic components except multiples of 3 are eliminated in the three-phase waveform, the 6th harmonics component is the dominant component in three-phase TFPM. In fact, its amplitude is three times of the 6th component of the single-phase machine. The other noticeable component is the 12th component which is not comparable to the 6th component. Consequently, the cogging torque of the three-phase TFPM is mainly determined by the 6th harmonics component. So, the cogging torque of the three-phase TFPM can be minimized by minimizing the 6th harmonic component. To do so, axial PM segmentation is adopted with a segmentation angle of φ_m. In this regard, Equation (4) should be solved. In this equation, the first term represents the torque component of the upper PMs and the second term represents it for the lower PMs.

$$T_{n=6} = a_6 \sin(6\theta_e) + a_6' \sin(6(\theta_e + \frac{40}{2}\varphi_m)) = 0 \tag{4}$$

By considering a_6 and a_6' as equal parameters and solving Equation (4), the possible values of shifting angles are determined as in Equation (5).

$$\varphi_m = 1.5 \times k \quad , \quad k = 1,3,5,\ldots \tag{5}$$

It is revealed from Equation (5) that ideally the 6th harmonic component of the cogging torque waveform is canceled if by choosing the shift angle as one of the following values; 1.5, 4.5, 7.5, etc.

To evaluate the validity of the above-mentioned solution, peak to peak value of the three phase cogging torque waveform is calculated using FEA at different shift angles. The results of such analysis are extracted and presented in Figure 5. It is noticed that by increasing the shifting angle, the torque reduces and reaches the minimum value at the angle of 1.5 degrees. At this point, the 6th harmonic component has the minimum value leading to the minimum cogging torque. By further increasing the shift angle, the cogging torque increases and reaches the peak value at the angle of 3 degrees. According to Equation (4), at this angle, the 6th harmonic component of cogging torque for top and bottom PMs are in the same phase angle which results in peak cogging torque value. By increasing the segmentation angle until 4.5 degrees, the cogging torque reduces to its minimum value. The cogging torque waveform of the TFPM is presented in Figure 6 for 0 and a 1.5-degree segmentation angle. The comparison shows that the cogging torque decreases by more than 85%. Figure 7 indicates harmonic components of the single-phase TFPM at 0 and 1.5-degree segmentation angles. The 6th harmonic component which mostly determines the three-phase cogging torque has decreased to 13% after the adoption of a 1.5-degree segmentation angle. Moreover, the 12th harmonic component has decreased by 20%.

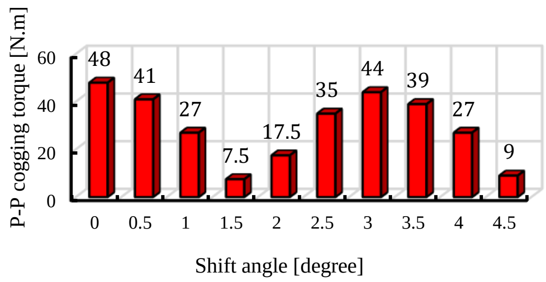


Figure 5. Peak to peak cogging torque values for three-phase TFPM versus axial segmentation angle

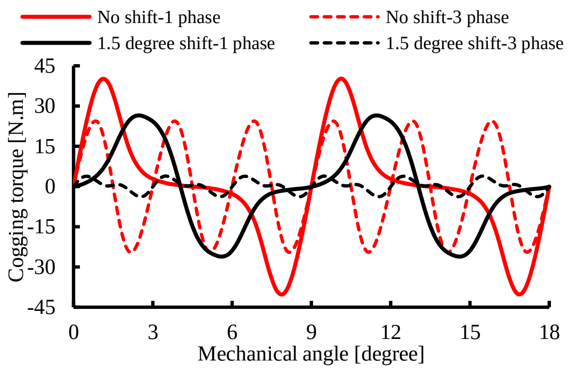


Figure 6. Single-phase and three-phase cogging torque waveforms of the TFPM generator at two segmentation angles

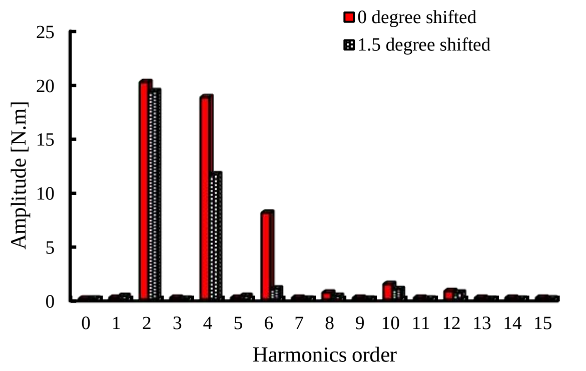


Figure 7. Harmonic components of single-phase TFPM at two segmentation angles

In this section, the PM segmentation concept was explained in TFPM topologies. Related formulations were extracted and optimum shift angles are formulized to achieve a minimum cogging torque value. By selecting a 1.5-degree segmentation angle 6th component of the cogging torque which is the dominant component of the cogging torque waveform reduces dramatically. So, it was revealed that after the adoption of this method the peak to peak value of three-phase cogging torque dropped from 48 N.m to 7.5 N.m. In the next section, the design and modeling algorithm of the TFPM generator is presented.

3. PROPOSING THE DESIGN ALGORITHM AND NON-LINEAR 3-D MEC

3. 1. Design Algorithm

In Figure 8, the proposed design and modeling algorithm is presented. At the first step, machine rated parameters such as nominal power and terminal voltage are inserted. Next, machine design parameters such as electrical and magnetic loadings, air gap clearance, PM characteristics are set. Then, the initial assumptions for power factor, efficiency, and iron part potential drops are inserted. Afterward, the rotor pole arc length and coil turn number per phase are determined. Electrical loading of the TFPM generator is calculated as in Equation (6);

$$A_m = \frac{\sqrt{2} \times N \times I_a}{2 \times \tau} \tag{6}$$

Here, I_a is rms value of the rated armature current, N is the number of coil turn per phase and τ is the rotor pole arc length which is driven from Equation (7). By defining the pole arc length to pole axial length ratio (τ/L_m) as machine aspect ratio (β):

$$\tau = \sqrt[3]{\frac{E_f \times \beta \times I_a}{2 \times f \times p \times k_f \times B_{mg} \times A_m}} \tag{7}$$

where B_{mg} is the maximum flux density in the middle of the air gap, E_f is rms value of phase Back-emf, k_f is the

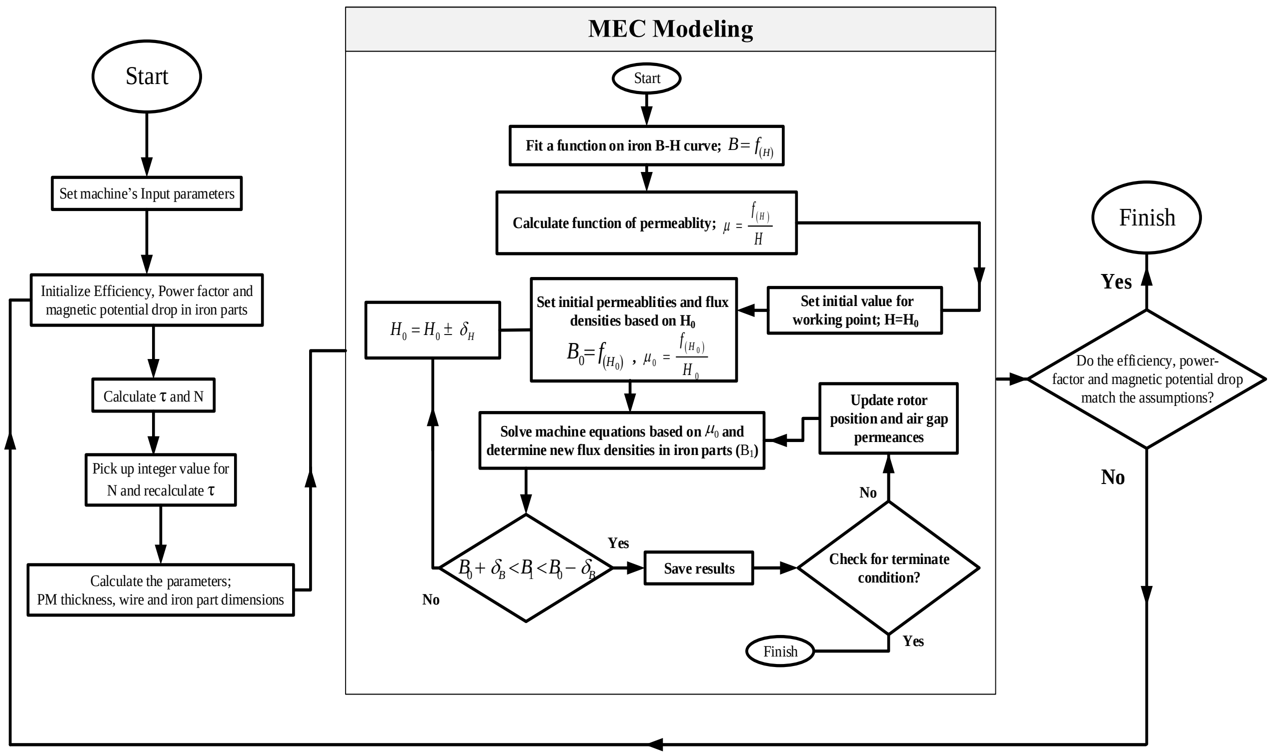


Figure 8. Proposed design algorithm of the TFPM coupled with non-linear MEC method

flux distribution factor, and f is the rated electrical frequency. After calculation of pole arc length, the number of coil turns is calculated using Equation (6). By rounding the value of N, the pole pitch is recalculated as in Equation (8).

$$\tau = \sqrt{\frac{E_f \times \beta}{\sqrt{2} \times f \times p \times k_f \times B_{mg} \times N}} \tag{8}$$

Based on Ampere's law in the main flux path of the magnetic circuit (Figure 1), the PM thickness is calculated as in Equation (9).

$$H_m = \frac{\mu_r B_{mg} \times g + 0.25\mu_0\mu_r \times V_{Fe}}{B_r - B_{mg}} \tag{9}$$

where, μ_r, B_r, g, and V_{Fe} are the PM relative permeability, PM residual flux density, air gap clearance, and the iron parts magnetic potential drop in one pole pair, respectively. Next, wire size, iron part dimensions such as the U-I thickness, rotor yoke thickness and window dimensions are calculated based on the initial inputs. At this step, an initial design for a TFPM is obtained. Now, the designed TFPM is analyzed using non-linear dynamic magnetic equivalent circuit method (MEC) which will be elaborated in the next subsection. After analyzing the TFPM using the MEC method, the algorithm checks if the assumed values of the efficiency, the power factor, and iron part potential drops are converged with the MEC results. If the assumed parameters converge with the MEC results, the output design is valid and the design procedure ends, otherwise, the assumed parameters are updated and the procedure goes on until the convergence results.

3. 2. Three Dimensional MEC Modeling In subsection 3.1, a detailed design algorithm of the TFPM generator was discussed. The proposed algorithm uses the non-linear MEC method to check the convergence of the design initial assumptions and the analysis results. So, accurate MEC modeling results in a more accurate design. Since, this study discusses the axial segmentation effect of the PMs on the cogging torque, the effect of two-step PM segmentation is included in the MEC method. In the proposed MEC method, effects of iron saturation, rotor movement, and fringing flux are also considered. The proposed MEC model will be used in the optimization procedure in the next section.

Two cross-sectional views of the TFPM reluctance network is presented in Figures 9(a) and 9(b). The front view of the proposed TFPM reluctance network is depicted in Figure 9(a) where the variable air gap reluctances are excluded. In this view, two-step PM segmentation is not visible while it is noticeable in the side view of Figure 9 (b). In this view, the I-shaped core is positioned behind the U-shaped core, consequently, the reluctance of this part is sketched using dotted lines (Figure 9 (b)).

3. 2. 1. General Equations Here, general equations of the MEC method are categorized as the followings:

PM Model: Each PM is modeled by a flux source shunted by a reluctance which represents the internal leakage of the PM. Corresponding values of the flux source and internal leakage reluctance are given in Equations (10) and (11), respectively.

$$F_1 = B_r \times L_m \times \tau_m \tag{10}$$

$$R_3 = \frac{d_m}{\mu_0 \mu_r \times L_m \times \tau_m} \tag{11}$$

Armature MMF and Back-EMF: Armature reaction is modeled by MMF source. The flux linkage is the total

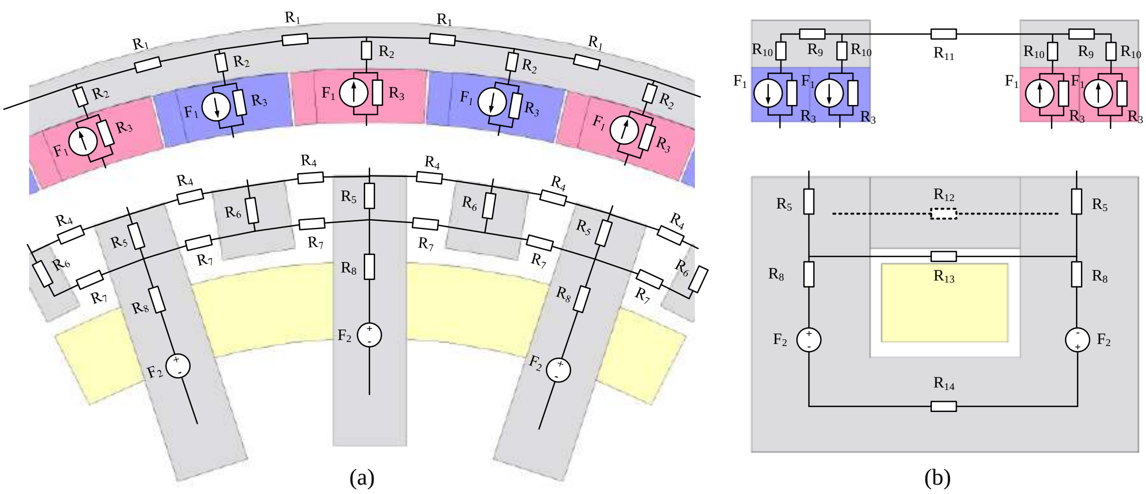


Figure 9. Reluctance network of the TFPM generator from two views, (a) front view, and (b) side view

flux which is in the same branch as the MMF source which determines the induced voltage in the coil as in Equation (12).

$$\begin{cases} F_2 = N \times I \\ \psi = N \times \phi \\ e = \dfrac{\partial \psi}{\partial t} \end{cases} \tag{12}$$

Torque Equation: According to literature [23, 24] the torque production is due to the change in the air gap permeance and the MMF drop on those permeances. So, for one phase of the TFPM the electromagnetic torque is calculated by Equation (13):

$$T = \sum_{i=1}^{n} U_i \times \frac{\partial G_i}{\partial \theta} \tag{13}$$

Here, G_i is the i[th] air gap permeance, U_i is MMF drop on it, θ is the rotor position, and n is the number of air gap permeances in each phase.

3. 2. 2. Variable Air Gap Permeance Model

The permeances between rotor poles and stator teeth are referred to as variable air gap permeances and are responsible for the energy conversion process. The air gap permeance modeling is based on the proposed model in literature [24]. The permeance model between one rotor pole (r_1) and one stator tooth (s_1) is presented in Figure 10. When the distance between the rotor pole and the stator tooth is as one pole pitch, the permeance is zero (Figure 10 (a)). By rotor movement and getting closer to stator teeth, the permeance increases until a complete overlap. This position is illustrated in Figure 10(b). By further movements of the rotor, the overlap is reduced and the permeance drops. It reaches zero when the position difference gets more than one pole pitch as presented in Figure 10(c). The transition of the permeance is assumed sinusoidal between G_{max} and zero. Consequently, the permeance between r_1 and s_1 is modeled as in Figure 10(d). The relevant parameters are defined as in Equation (14).

$$\left\{ a = \frac{2\pi}{p},\ b = \frac{|\tau_m - T_u|}{D_g} \right. \tag{14}$$

The value of variable air gap permeance is composed of two main components. The major component is caused by the overlapped area of the r_1 and s_1. Here, the flux path is straight and is illustrated in Figure 10 (e). The second component is due to the fringing flux and is presented in Figure 10(f), consequently, the maximum value of variable air gap permeance is calculated as in Equation (15) which the first term represents overlapping permeance and the second term represents the fringing permeance.

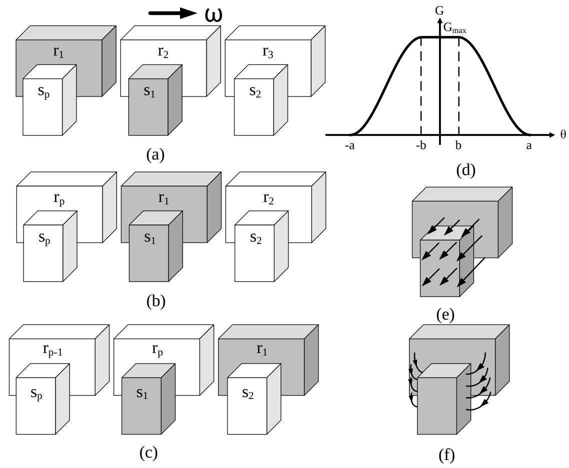


Figure 10. The variable air gap permeance between rotor pole (r_1) and stator tooth (s_1) at different angles, (a) angle of –a, (b) angle of zero, (c) angle of a, (d) Permeance model between r_1 and s_1, (e) Overlapping term of variable air gap permeance, and (f) Fringing term of variable air gap permeance

$$G_{max} = \left[\frac{\mu_0 L_m \times \min(\tau_m, T_u)}{g} \right] + \left[\frac{4\mu_0 L_m}{\pi} \times Ln(1 + \frac{\pi|\tau_m - T_u|}{4g}) \right] \tag{15}$$

In the case of PM segmentation, each PM is axially divided into two PMs which the lower PM is shifted by the angle of θ_s. Consequently, if the permeance function between the top PM and the stator tooth (s_1) is $G_{(\theta)}$, then the permeance function between the bottom PM and the s1 is $G_{(\theta-\theta_s)}$.

3. 2. 3. Iron Saturation Effect To take the effect of iron saturation into account in the MEC model, the steel sheet B-H curve is split into four regions and a function is fitted per region. So, the steel B-H curve is estimated with a high level of precision using a piecewise function ($B=f_{(H)}$). Consequently, the permeability of the steel can be described by (16);

$$\mu = \frac{f(H)}{H} \tag{16}$$

At first, a random working point on the steel B-H curve is assumed for iron parts. Next, the corresponding reluctances are calculated for these parts based on the initial working point. Next, the reluctance network of the TFPM is solved and the magnetic flux densities of iron parts along with actual working point on steel B-H curve are obtained. The resulted working point is compared with the initial working point and if the difference is smaller than a specific value (δ_B), the results are saved and the next time step starts. Otherwise, the assumed working point is not valid, so the working point is changed till the difference between the initial and the resulted working points is smaller than the accepted error

(δ_B). Based on the aforementioned discussion, the algorithm of the non-linear dynamic MEC model is given in Figure 8.

3. 2. 4. MEC Results The torque and voltage waveforms of the 2kW TFPM are extracted using the proposed MEC method and compared with the FEM results in terms of precision and computational time. The no-load induced voltage waveform of the TFPM is calculated by both the MEC method and FEM and depicted in Figure 11. The result of the MEC method shows acceptable accuracy in calculating the Back-emf, since the point to point error is lower than 4%. Similarly, the cogging torque and the rated load torque are presented in Figure 12 and Figure 13, respectively. The comparison shows that the MEC method calculates the machine torque by less than 6% error in the worst case. The computational time of either method is also measured which reveals that the MEC method nearly 10 times faster than the FEM.

In this section, the design and modeling procedure of the TFPM generator was discussed. A comprehensive algorithm was proposed where the MEC method was adopted to increase the accuracy of the design procedure. The detail of the proposed MEC method was elaborated when modeling the effects of iron saturation, PM segmentation, and rotor movement were discussed. The results of the MEC were compared with the FEM results which both accuracy and computational speed of the proposed MEC method were proven. In the next section, the Taguchi optimization method is used to improve the torque characteristics of the designed TFPM generator.

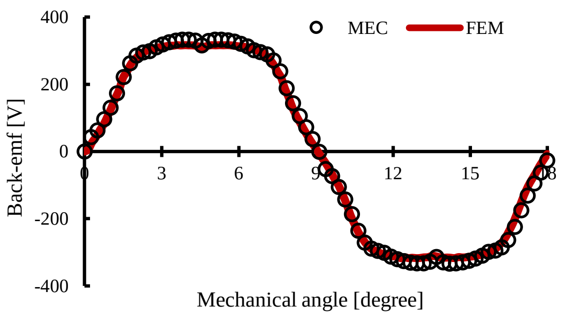


Figure 11. Back-emf waveforms of the TFPM generator at no-load condition resulted from the MEC method and FEM

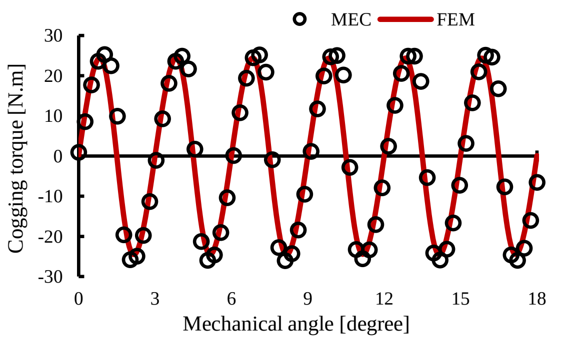


Figure 12. Cogging torque waveforms of the three-phase TFPM generator resulted by MEC method and FEM

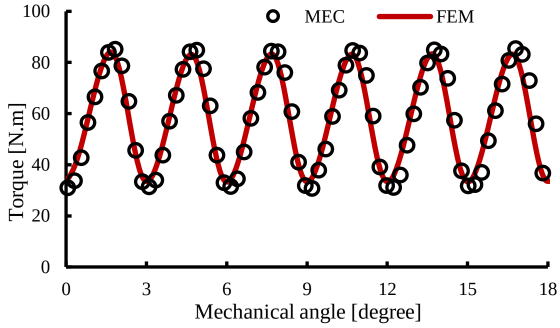


Figure 13. Electromagnetic torque waveform of the TFPM generator at rated load resulted by MEC and FEM

4. OPTIMIZATION OF TORQUE PROFILE

Until now, the PM segmentation procedure is discussed and the optimum shift angle between the upper and lower PM segments is formulated. Accordingly, the Optimum shift angle is fixed at 1.5 degrees (Equation (5)), and the sixth harmonic component which is the dominant component is minimized. Consequently, the peak cogging torque value of the TFPM has dropped from 24 N.m to 3.8 N.m after PM segmentation. To minimize the cogging torque as much as possible, the optimization method is used to optimize the machine parameters while the segmentation shift angle is fixed at its optimum value.

4. 1. Taguchi Optimization Method In this section, torque profiles of the 2 kW TFPM are optimized to improve its performance as a DDWT generator. Taguchi optimization method is selected as it is a popular and effective optimization method [25, 26]. Taguchi optimization method is one of the designs of experiments (DOE) methods which its accuracy, computational speed, and robustness have been proven. In an optimization problem, if there are 3 variables each at 5 levels, the full factorial design requires 5^3=125 experiments while the Taguchi method offers to perform only 25 experiments which leads to a considerable amount of computational effort.

4. 2. Optimization Procedure Aiming to minimize the cogging torque of the TFPM generator, 3 parameters each at 5 levels are chosen as optimization parameters. These variable parameters are as listed as following: stator pole shoe width (T_u), PM pole width to pole pitch ratio (α), and PM thickness (d_m).

Due to choosing 3 variables at 5 levels, the L_{25} orthogonal array is selected. At each experiment, cogging torque and the average torque are calculated and analyzed. In this study, achieving a minimum value for the ratio of cogging torque to average electromagnetic torque is the goal of the optimization.

4. 3. Optimization Results Figure 14 and Figure 15 represent the mean effect of variables on cogging torque and average torque, respectively. It is noticed that the first level of T_u, the third level of α, and the fifth level of d_m are the best combination which minimizes the cogging torque. From Figure 15, it is concluded that the average torque increases by increasing the variable levels. Hence, the combination of the fifth level of variables leads to the maximum value of average torque. Mean effect of the variables on cogging to average torque ratio which is the optimization goal is presented in Figure 16. As it can be noticed, the optimization function is minimized by choosing the 1st level of T_u, 3rd level of α, and 5th level of d_m. By selecting the aforementioned combination, and analyzing the final TFPM generator, the cogging torque waveforms are extracted and presented in Figure 17. The cogging torque of the final TFPM is reduced by 85% with respect to the non-optimized TFPM and 97% with respect to the initial design which is a dramatic reduction in generator cogging torque. The electromagnetic torque of the TFPM is also extracted at rated current compared with the non-optimized and initial designs (Figure 18). The comparison shows that the torque ripple of the TFPM after optimization has improved significantly. However, the average torque decreased by 8% with respect to the initial design and 5% with respect to the non-optimized TFPM. The mentioned decrease in the machine average torque is the most important side effect of the PM segmentation. The second side effect of the method is increased manufacturing complexity because of the two-step skew technique. In this way, instead of placing one PM per pole (conventional design), two PMs per pole are required where increases the manufacturing complexity of the machine. Nevertheless, losing 8% of the average torque and relatively higher complexity are tolerable considering a 97% decrease in the machine cogging torque which is a significant achievement compared to the existing methods in the literature. Finally, a cross-sectional view of the TFPM after 1.5 degrees PM segment shift and optimization procedure is presented in Figure 19. Contour and vector plots of magnetic flux density distribution in the TFPM along with the mesh plot of the TFPM are presented in Figure 19.

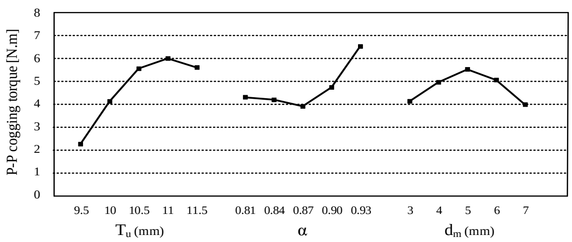


Figure 14. Mean effect of optimization variables on cogging torque of the three-phase TFPM generator

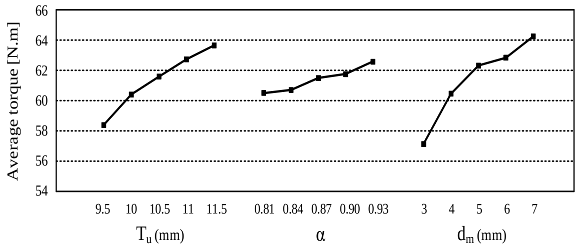


Figure 15. Mean effect of optimization variables on average torque of the three-phase TFPM generator

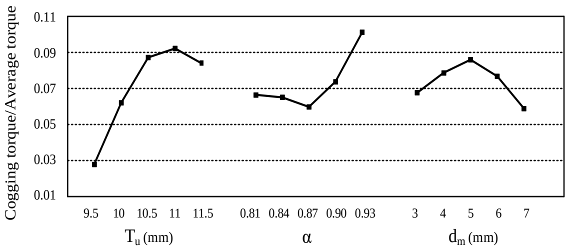


Figure 16. Mean effect of optimization variables on the ratio of TFPM cogging torque to average torque

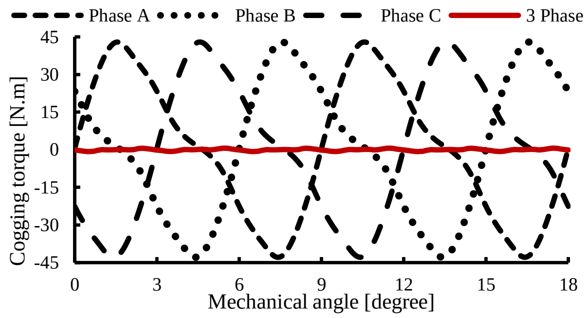


Figure 17. Single-phase and three-phase cogging torque waveforms of the optimized TFPM generator

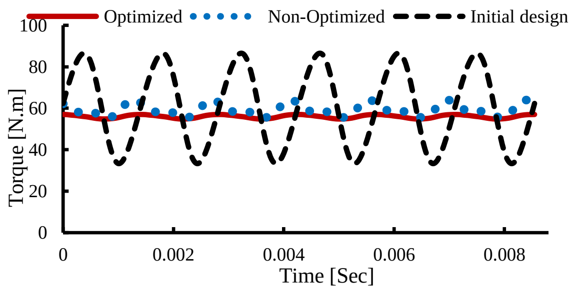


Figure 18. Electromagnetic torque of the three-phase TFPM generators at rated current

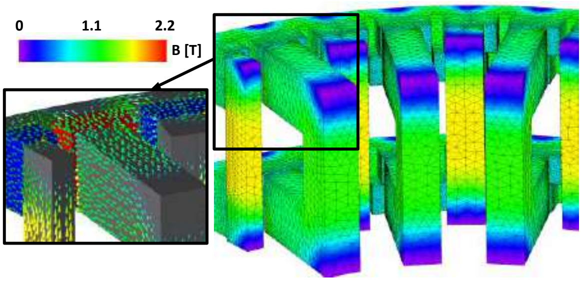


Figure 19. The cross-sectional view of the final TFPM generator along with vector and contour plots of magnetic flux density distribution

5. CONCLUSION

In this study, the optimum design of the TFPM generator for direct driven wind turbine application was addressed. Two steps axial PM segmentation concept in TFPMs was explained and the optimum shift angle was formulated. A non-linear dynamic MEC method was developed and a MEC based design algorithm for a two kW TFPM was presented. Results of the MEC were compared with the FEM results. The comparison showed that by replacing the MEC with the FEM, noticeable processing time was saved where acceptable level of accuracy was also resulted.

By adopting two step axial PM segmentation at the optimum segmentation angle, the cogging torque of the TFPM generator decreased by 85% with respect to the non-segmented generator. Taguchi optimization method was applied to minimize the ratio of cogging torque to average electromagnetic torque. The peak value of cogging torque was reduced from 24 N.m to 0.7 N.m after performing both the axial segmentation technique and Taguchi optimization method on the initial TFPM generator. Nevertheless, the generator torque rating was reduced by 8%. But, by considering the significant reduction in cogging torque (97%), losing 8% of generator rated torque can be easily tolerated.

6. REFERENCES

1. Sang-Yong Jung, Hochang Jung, Sung-Chin Hahn, Hyun-Kyo Jung, and Cheol-Gyun Lee, "Optimal Design of Direct-Driven PM Wind Generator for Maximum Annual Energy Production", ***IEEE Transactions on Magnetics***, Vol. 44, No. 6, (2008), 1062–1065. doi:10.1109/TMAG.2007.916250

2. McDonald, A., and Bhuiyan, N. A., "On the Optimization of Generators for Offshore Direct Drive Wind Turbines", ***IEEE Transactions on Energy Conversion***, Vol. 32, No. 1, (2017), 348–358. doi:10.1109/TEC.2016.2624219

3. Ghaheri, A., Mohammadi Ajamloo, A., Torkaman, H., and Afjei, E., "Design, modelling and optimisation of a slot-less axial flux permanent magnet generator for direct-drive wind turbine application", ***IET Electric Power Applications***, Vol. 14, No. 8, (2020), 1327–1338. doi:10.1049/iet-epa.2019.0385

4. Bhuiyan, N. A., and McDonald, A., "Optimization of Offshore Direct Drive Wind Turbine Generators With Consideration of Permanent Magnet Grade and Temperature", ***IEEE Transactions on Energy Conversion***, Vol. 34, No. 2, (2019), 1105–1114. doi:10.1109/TEC.2018.2879442

5. Nasiri-Zarandi, R., Ghaheri, A., and Abbaszadeh, K., "Thermal Modeling and Analysis of a Novel Transverse Flux HAPM Generator for Small-Scale Wind Turbine Application", ***IEEE Transactions on Energy Conversion***, Vol. 35, No. 1, (2020), 445–453. doi:10.1109/TEC.2019.2936683

6. Moradi CheshmehBeigi, H., "Design, Optimization and FEM Analysis of a Surface-Mounted Permanent-magnet Brushless DC Motor", ***International Journal of Engineering***, Vol. 31, No. 2, (2018), 339–345. doi:10.5829/ije.2018.31.02b.19

7. Zahabi, M., Ardeshir, G., and Ale Ahmad, A., "Influence of DC-Link Voltage on Commutation Torque Ripple of Brushless DC Motors with Two-Segment Pulse-width Modulation Control Method", ***International Journal of Engineering - Transaction B: Applications***, Vol. 31, No. 2, (2018), 307–314. doi:10.5829/ije.2018.31.02b.15

8. Torkaman, H., Ghaheri, A., and Keyhani, A., "Design of Rotor Excited Axial Flux-Switching Permanent Magnet Machine", ***IEEE Transactions on Energy Conversion***, Vol. 33, No. 3, (2018), 1175–1183. doi:10.1109/TEC.2018.2807804

9. Husain, T., Hasan, I., Sozer, Y., Husain, I., and Muljadi, E., "Design Considerations of a Transverse Flux Machine for Direct-Drive Wind Turbine Applications", ***IEEE Transactions on Industry Applications***, Vol. 54, No. 4, (2018), 3604–3615. doi:10.1109/TIA.2018.2814979

10. Ajamloo, A. M., Ghaheri, A., and Nasiri-Zarandi, R., "Design and Optimization of a New TFPM Generator with Improved Torque Profile", 2019 International Power System Conference (PSC), IEEE, (2019), 106–112. doi:10.1109/PSC49016.2019.9081559

11. Ueda, Y., and Takahashi, H., "Transverse-Flux Motor Design With Skewed and Unequally Distributed Armature Cores for Reducing Cogging Torque", ***IEEE Transactions on Magnetics***, Vol. 53, No. 11, (2017), 1–5. doi:10.1109/TMAG.2017.2703087

12. Wang, Q., Zhao, B., Zhao, H., Li, Y., and Zou, J., "Optimal Design of Tubular Transverse Flux Motors With Low Cogging Forces for Direct Drive Applications", ***IEEE Transactions on Applied Superconductivity***, Vol. 26, No. 7, (2016), 1–5. doi:10.1109/TASC.2016.2600104

13. Washington, J. G., Atkinson, G. J., and Baker, N. J., "Reduction of Cogging Torque and EMF Harmonics in Modulated Pole Machines", ***IEEE Transactions on Energy Conversion***, Vol. 31, No. 2, (2016), 759–768. doi:10.1109/TEC.2016.2520200

14. Nasiri-Zarandi, R., Ajamloo, A. M., and Abbaszadeh, K., "Design Optimization of a Transverse Flux Halbach-Array PM Generator for Direct Drive Wind Turbines", ***IEEE Transactions on Energy Conversion***, Vol. 35, No. 3, (2020), 1485–1493. doi:10.1109/TEC.2020.2975259

15. Nasiri-Zarandi, R., Ghaheri, A., and Abbaszadeh, K., "Cogging Torque Reduction in U-Core TFPM Generator Using Different Halbach-Array Structures", 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), IEEE, (2018), 1153–1158. doi:10.1109/SPEEDAM.2018.8445328

16. Liu, C., Zhu, J., Wang, Y., Lei, G., and Guo, Y., "Cogging Torque Minimization of SMC PM Transverse Flux Machines Using Shifted and Unequal-Width Stator Teeth", ***IEEE Transactions on Applied Superconductivity***, Vol. 26, No. 4, (2016), 1–4. doi:10.1109/TASC.2016.2543959

17. Dobzhanskyi, O., Gouws, R., and Amiri, E., "Analysis of PM Transverse-Flux Outer Rotor Machines With Different Configuration", ***IEEE Transactions on Industry Applications***, Vol. 53, No. 5, (2017), 4260–4268.

doi:10.1109/TIA.2017.2696901

18. Gieras, J. F., "Performance Characteristics of a Transverse Flux Generator", IEEE International Conference on Electric Machines and Drives, 2005, IEEE, (2005), 1293–1299. doi:10.1109/IEMDC.2005.195889
19. Azari, M. N., "Performance Improvement of a Slotted Solid Rotor Induction Motor with High Temperature Superconder Coating", ***International Journal of Engineering - Transaction B: Applications***, Vol. 32, No. 5, (2019), 693–700. doi:10.5829/ije.2019.32.05b.11
20. Fekri, H., Shamsi-Nejad, M., and Hasheminejad, S., "Performance Analysis of a Novel Three-phase Axial Flux Switching Permanent Magnet Generator with Overlapping Concentrated Winding", ***International Journal of Engineering - Transaction B: Applications***, Vol. 32, No. 2, (2019), 286–295. doi:10.5829/ije.2019.32.02b.14
21. Ostovic, V., "Dynamics of saturated electric machines", Springer Science & Business Media, (2012).
22. Nasiri-Zarandi, R., Ajamloo, A. M., and Abbaszadeh, K., "Proposing the Output Equations and 3-D MEC Modeling for U - Core TFPM Generators", 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), IEEE, (2018), 292–297. doi:10.1109/SPEEDAM.2018.8445324
23. Sudhoff, S.D., "Power magnetic devices: A multi-objective design approach", John Wiley & Sons, (2014).
24. Mohammadi Ajamloo, A., Ghaheri, A., Shirzad, H., and Afjei, E., "Non-linear analytical modelling and optimisation of a 12/8 rotor excited flux-switching machine", ***IET Electric Power Applications***, Vol. 14, No. 9, (2020), 1592–1603. doi:10.1049/iet-epa.2019.0732
25. Ghorbani, S., and Reaz Kashyzadeh, K., "Taguchi Approach and Response Surface Analysis for Design of a High-performance Single-walled Carbon Nanotube Bundle Interconnects in a Full Adder", ***International Journal of Engineering - Transaction B: Applications***, Vol. 33, No. 8, (2020), 1598–1607. doi:10.5829/ije.2020.33.08b.18
26. Mohammadi Ajamloo, A., Abbaszadeh, K., and Nasiri-Zarandi, R., "A Novel Transverse Flux Permanent Magnet Generator for Small-Scale Direct Drive Wind Turbine Application: Design and Analysis", ***Scientia Iranica***, (2019). doi:10.24200/sci.2019.53145.3078

Persian Abstract

چکیده

ماشین‌های آهنربای دایم شارمتقاطع گونه‌ای از ماشین های سنکرون هستند که از چگالی گشتاور زیادی برخوردارند و قابلیت جای دادن تعداد قطب های زیادی را دارا هستند. این دو خصوصیت منجر می‌شود تا ماشین‌های شارمتقاطع گزینه مناسبی برای کاربردهای کم سرعت با چگالی گشتاور بالا از جمله توربین بادی وصل مستقیم باشند. علیرغم مزایای نام برده، ماشین‌های شارمتقاطع به طور ذاتی از گشتاور دندانه زیاد رنج می‌برند که ویژگی نگران‌کننده‌ای در توربین‌های بادی وصل مستقیم می‌باشد. هدف این مقاله، کمینه کردن گشتاور دندانه این ماشین‌ها با استفاده از روش تکه کردن آهنربا در جهت محوری می‌باشد. مفهوم این روش پیشنهادی با استفاده از روش‌های تحلیلی بیان شده و میزان اختلاف زاویه بهینه آهنرباها در یک فرمول جامع استخراج شده است. روش مدار معادل مغناطیسی غیرخطی در این مقاله به کار گرفته شده است که در آن اثر تکه کردن آهنربا، عکس العمل آرمیچر، گردش روتور و اشباع هسته آهن به دقت مدل شده است. همچنین، نتایج تحلیل مدار معادل مغناطیسی با نتایج روش المان محدود از حیث دقت و زمان شبیه‌سازی مقایسه شده است. نتیجه مقایسه حاکی از آن است که روش مدار معادل مغناطیسی پیشنهادی زمان شبیه‌سازی را تا 10 برابر نسبت به روش المان محدود کاهش داده و از میزان دقت قابل قبولی برخوردار می‌باشد. سپس روش بهینه‌سازی تاگوچی به عنوان یکی از روش‌های بهینه‌سازی پرسرعت بکار گرفته شده است تا مشخصات گشتاور ژنراتور شار متقاطع را بهبود ببخشد. نتایج بهینه‌سازی حاکی از آن است که با استفاده از روش پیشنهادی این مطالعه و انجام فرآیند بهینه‌سازی، گشتاور دندانه ژنراتور 97 درصد کاهش پیدا کرده است درحالی که گشتاور میانگین حدود 8 درصد افت نموده است که با توجه به افت قابل توجه گشتاور دندانه، می‌تواند به عنوان یک اثر جانبی، قابل تحمل باشد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 919-927

International Journal of Engineering

Journal Homepage: www.ije.ir

Application of Artificial Neural Network and Multi-magnetic NDE Methods to Determine Mechanical Properties of Plain Carbon Steels Subjected to Tempering Treatment

I. Ahadi Akhlaghi[a], S. Kahrobaee[b], M. Sekhavat[c], H. Norouzi Sahraei[d], F. Akhlaghi Modiri[e]

[a] *Department of Electrical and Bioelectric Engineering, Sadjad University of Technology, Mashhad, Iran*
[b,c] *Department of Mechanical and Materials Engineering, Sadjad University of Technology, Mashhad, Iran*
[a,b,d,e] *Center of Nondestructive Evaluation (CNDE), Sadjad University of Technology, Mashhad, Iran*

PAPER INFO

Paper history:
Received 12 July 2020
Received in revised form 02 January 2021
Accepted 14 January 2021

Keywords:
Mechanical Parameters
Magnetic Hysteresis Loop
Magnetic Flux Leakage
Eddy Current
Artificial Neural Network

ABSTRACT

The present paper shows the results of applying an artificial neural network to three non-destructive magnetic methods including magnetic hysteresis loop (MHL), eddy current (EC), and magnetic flux leakage (MFL) techniques to determine mechanical features of plain carbon steels with unknown carbon contents subjected to tempering treatment. To simultaneously evaluate the effects of carbon content and microstructure on the magnetic and mechanical properties, four grades of hypoeutectoid steel samples containing 0.30, 0.46, 0.54, and 0.71 wt.% carbon were austenitized in the range of 830-925 °C and then subjected to quench-tempering treatments at 200, 300, 400, 500 and 600 °C. In the next step, mechanical properties including tensile strength, elongation, and hardness were measured using tensile and hardness tests, respectively. Finally, to study the electromagnetic parameters, MHL, MFL and EC non-destructive electromagnetic tests were applied to the heat-treated samples and their outputs were fed to a generalized neural network designed in this work. The results revealed that using a proper combination of electromagnetic parameters as the ANN input for each mechanical parameter enables us to determine the hardness, UTS and elongation of hypoeutectic carbon steel parts after tempering treatment with high accuracy.

***doi**: 10.5829/ije.2021.34.04a.18*

1. INTRODUCTION

In recent years, there is a high demand for reliable and quick methods to be used for nondestructive characterization of materials. Electromagnetic techniques such as eddy current (EC) and magnetic hysteresis loop (MHL) measurements with a high sensitivity to microstructural changes have the potential to be used as alternatives to the traditional methods for the examination of microstructure and mechanical properties of steel parts [1]. Recent investigations in this field have been focused on microstructural changes due to the various heat-treating processes. For instance, pearlite/ferrite fraction [2], prior austenite grain size [3], microstructural changes during recovery process [4], the thickness of induction [5-7], carburized [8], and decarburized [9-11] layers in mild carbon steels, retained austenite fraction [12], precipitation of alloy carbides [13] and secondary hardening occurrence [14] in tool steels as well as variations in mechanical properties of API X65 [15], powder metallurgical [16] and Hadfield [17] steels have been characterized using EC and MHL methods.

Besides, the nondestructive magnetic flux leakage (MFL) method is a fast and easy technique to determine flaws in ferromagnetic materials and has widely used to characterize and determine hidden corrosion occurred in the gas ferromagnetic pipelines [18, 19], rail tracks [20] and wire ropes in suspension bridges [21]. In this method, Hall effect sensors sense the leakage flux in the area near the defect.

*Corresponding Author Institutional Email: i_a_akhlaghi@sadjad.ac.ir (I. Ahadi Akhlaghi)

Please cite this article as: I. Ahadi Akhlaghi, S. Kahrobaee, M. Sekhavat, H. Norouzi Sahraei, F. Akhlaghi Modiri, Application of artificial neural network and multi- magnetic NDE methods to determine mechanical properties of plain carbon steels subjected to tempering treatment, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 919-927

The proposed method in this paper takes advantage of artificial neural networks (ANN) to combine the results of the MHL, EC, and MFL electromagnetic subsystems described before. ANNs are inspired by the neural structure of the human brain and could be considered as a Machine Learning (ML) method. Clustering, regression, and classification problems could be solved using a well-trained ANN. Training an ANN consists of some standard mathematical algorithms by which the parameters of the network are determined and the trained network could then give correct answers to new inputs [22, 23].

Since ANNs may be more reliable, more accurate, and have better performance in many regression and classification problems than classic algorithms, they are widely used in practical applications of non-destructive evaluation methods. For example, in [14], after measuring four different electromagnetic properties of a given heat-treated H13 tool steel sample, Kahrobaee et al. have utilized a standard ANN (Generalized Regression Neural Network, GRNN) to estimate the hardness of the steel sample. They have reported that their proposed method could be used to non-destructively estimate the sample hardness with less than 1% RMS error if a training dataset with only 30 samples is used to train the network.

In the present paper, an application of a GRNN has been introduced to increase the prediction accuracy of the three MHL, MFL, and EC NDE electromagnetic methods in determining the mechanical properties of plain carbon steels with various carbon contents subjected to different tempering temperatures. In this methodology, a set of experimental data have been collected to obtain the initial database used to train and test the proposed ANN. The main novelties of the present paper are as follow:

- MHL, MFL, and EC outputs are combined to enhance the accuracy and reliability of determining the mechanical properties.
- NDE outputs obtained from the samples with known carbon contents and tempering temperatures are fed to the proposed ANN as intermediate variables that create a link from NDE outputs to the values of mechanical properties.
- The artificial neural network used in this research is a Generalized Regression Neural Network (GRNN). GRNNs could reportedly be trained much faster than standard Multilayer perceptron (MLP) neural networks and due to their high accuracy, they are widely used in regression, prediction, and classification problems.
- Apart from its high accuracy, using the GRNN as the estimator makes the proposed method more reliable and robust.

2. EXPERIMENTAL PROCEDURE

2. 1. Sample Preparation and Mechanical Tests

Four groups of plain carbon steel specimens (100×20×4 mm) containing various weight percentages of carbon (0.30, 0.46, 0.54, and 0.71) were austenitized for 50 min at 925°C, 900°C, 860°C, and 830°C, respectively. After water quenching all the samples, except for one in each group, the others were submitted to the tempering process at various temperatures (200 to 600 °C). To determine the mechanical features of the heat-treated samples as the ANN outputs, tensile tests were carried out according to the ASTM E8M tensile standard using Zwick testing machine. For each sample, an average value corresponds to three tensile specimens has been reported. For each sample, ultimate tensile strength (UTS) and elongation (E) were determined from its plotted engineering stress-strain graph. Moreover, hardness (Rockwell C) tests were performed at several points on each sample and the mean value has been reported as the third mechanical property.

2. 2. Magnetic Non-destructive Methods

After performing the mechanical tests to produce the outputs of the designed ANN, three nondestructive methods based on electromagnetic properties including MHL, MFL, and EC were carried out on the samples. The outputs of the NDT methods were utilized as the ANN inputs to evaluate their relationship to mechanical properties. Figure 1 demonstrates the schematic block diagram of MFL, MHL, and EC methods, whose outputs are used as the input of the proposed ANN.

In the MFL and MHL measurement systems, two 1500-turns in-series coils wound on two arms of a laminated Fe-Si U-shaped yoke are utilized as excitation coils to magnetize the sample. These coils are excited by a triangle waveform whose frequency and amplitude can be completely adjusted in a LabVIEW program. The waveform generated by the LabVIEW program, which is available from one of the analog ports of the Advantech PCI-1720U-AE DAC card, does not have enough power to magnetize the sample to saturation point. Therefore, a two-stage bi-polar and linear current amplifier, specifically designed and built for this and similar researches in our research laboratory, is used to amplify the excitation current.

In the MHL method, two other in-series 1000-turns coils wound on the former yoke are exploited to gauge the induced voltage (emf) in the samples. This voltage is sampled with the sampling rate of 500 Hz using an Advantech PCI-1714UL-BE ADC card and then is integrated over time and rescaled in a MATLAB program to give B-field (the magnetic flux density). To draw a hysteresis loop in the MHL method, the H-field (magnetic field strength) is also required. Since the field strength is completely proportional to the current given to excitation coils and the current amplifier used in this research is perfectly linear at the frequency of excitation (0.1 Hz), the excitation current and the waveform generated in the LabVIEW program have a linear relationship and therefore one could consider an

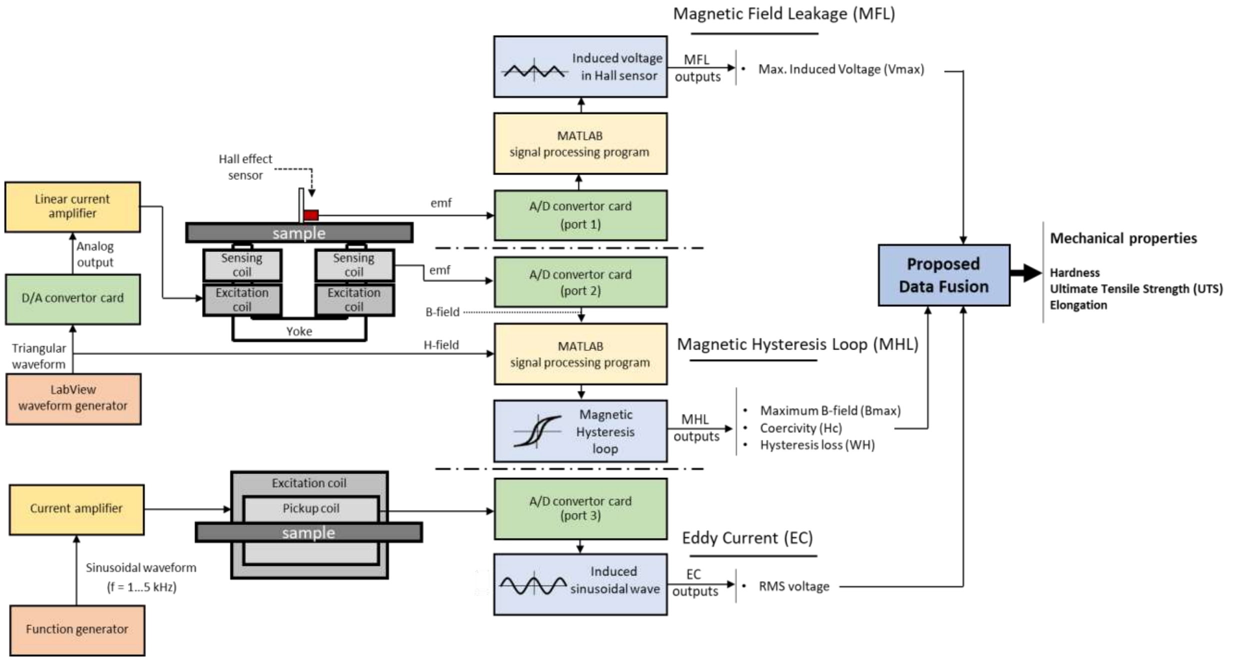


Figure 1. Schematic diagram of three nondestructive electromagnetic methods

appropriately scaled version of LabVIEW waveform as the magnetic field strength. Using both measured magnetic field strength and magnetic flux density, the BH curve (hysteresis loop) is drawn and three important magnetic parameters are extracted: Bmax (maximum B-field), Hc (coercivity), and WH (hysteresis loss).

In the MFL subsystem, a Hall effect sensor (HW-108C) is placed 1 mm over the sample to measure the magnetic flux leakage. This voltage is then given to another port of the ADC card and sampled with the sampling rate of 500 Hz. The magnetic flux leakage is a function of the sample parameters and its maximum value in a company with the output of MHL and EC subsystems could be used to extract mechanical properties of the sample.

In the EC method, two coaxially wound coils (excitation and pickup coils with respectively 400 and 800 turns) with a length of 30 mm, were used. A laboratory function generator is used to generate a 1-5 kHz sinusoidal waveform. After it is appropriately amplified, this waveform is fed to the excitation coil. The pickup coil voltage (which is mainly due to the eddy current inside the sample) is then given to the third analog port of the A/D conversion card and its root mean square (RMS) extracted using a MATLAB signal processing program, besides the results of MFL and MHL methods, is used as the inputs of the proposed DF.

2. 3. Structure of the Designed Artificial Neural Network In this research, a GRNN (Generalized Regression Neural Network) has been utilized to determine the mechanical properties of given plain carbon steel samples. GRNNs, which are based on RBF (Radial Basis Function) neural networks, were introduced in 1991 by D. F. Specht [24]. Figure 2 demonstrates the schematic block diagram of a typical GRNN [14]. In the input layer, one single neuron (neural cell) is dedicated to each input variable and in the radial basis layer, one radial-basis cell is dedicated to each training data point. Neurons of the third layer (output layer) are used to add the outputs of the former layer together and give out the network output. Since GRNNs can be easily trained much faster than other types of neural networks, they are increasingly used for solving different types of problems, like classification and regression problems.

3. RESULTS AND DISCUSSION

3. 1. Mechanical Properties Versus Tempering Temperature Figure 3 demonstrates the changes in mechanical properties for the samples containing various amounts of carbon subjected to different temperatures of the tempering process. Each point in the graphs is an average of three tests performed on three samples with constant conditions (carbon content and tempering temperature). As expected, tempering at higher temperatures reduces the values of hardness and UTS and enhances the elongation. A similar trend is observed for the steel samples containing various carbon contents: the higher amounts of hardness and UTS as well as the lower

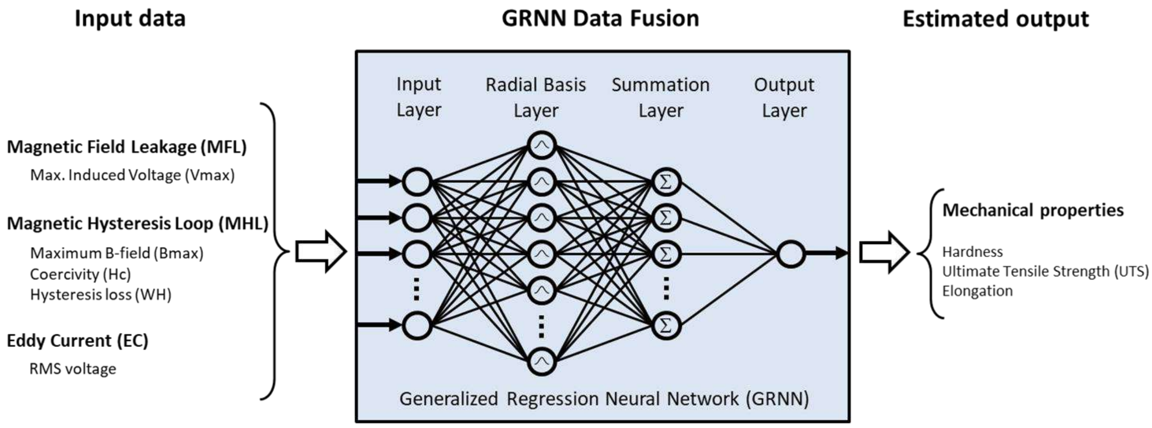


Figure 2. Schematic block diagram of a generalized neural network (GRNN).

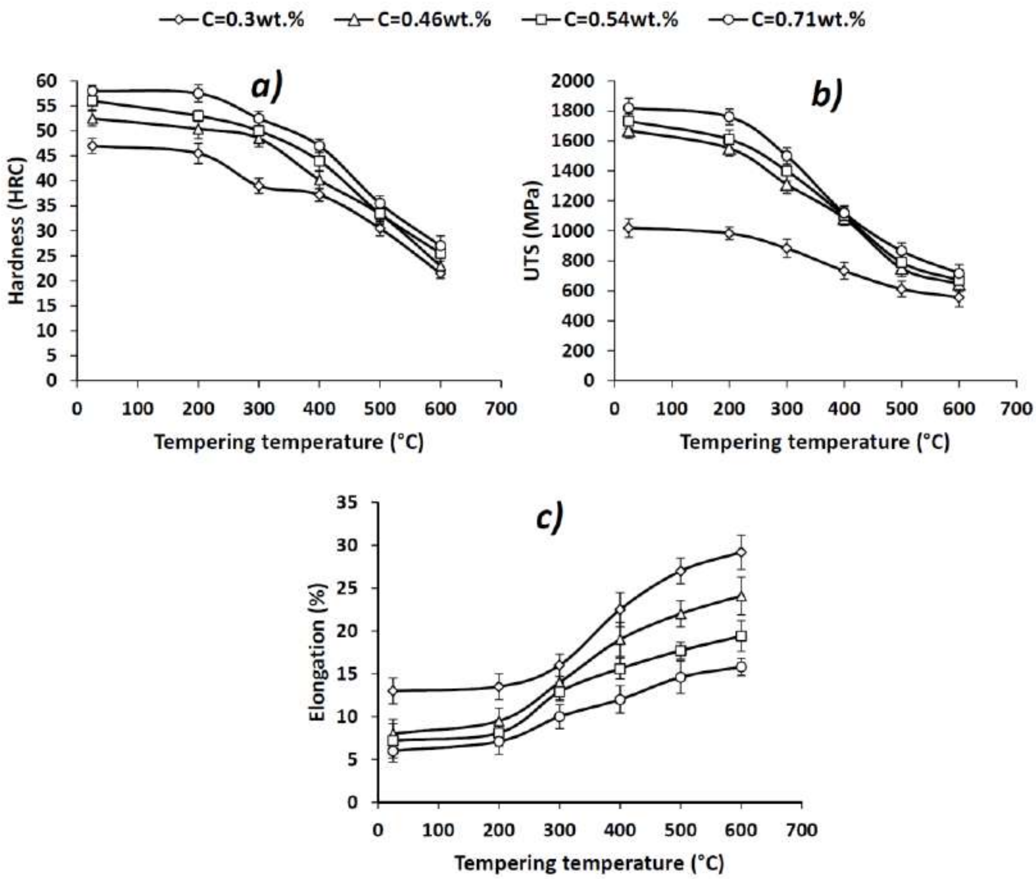


Figure 3. ANN outputs: mechanical properties including a) hardness, b) UTS, and c) elongation, as the function of tempering temperature for heat-treated samples containing different amounts of carbon

values of elongation have been obtained at the higher carbon contents.

3. 2. Magnetic NDE Parameters Versus Tempering Temperature

Figure 4 shows the output changes of the three NDE methods (Bmax, Hc, and WH for MHL, Vmax from MFL, and VRMS from EC) used in this work. The NDE tests were performed four times for all of the samples and the average values have been reported in this figure. The low and high values of respectively “Bmax and VRMS” and “Hc, WH and Vmax” for as-quenched samples in each group are mainly attributed to

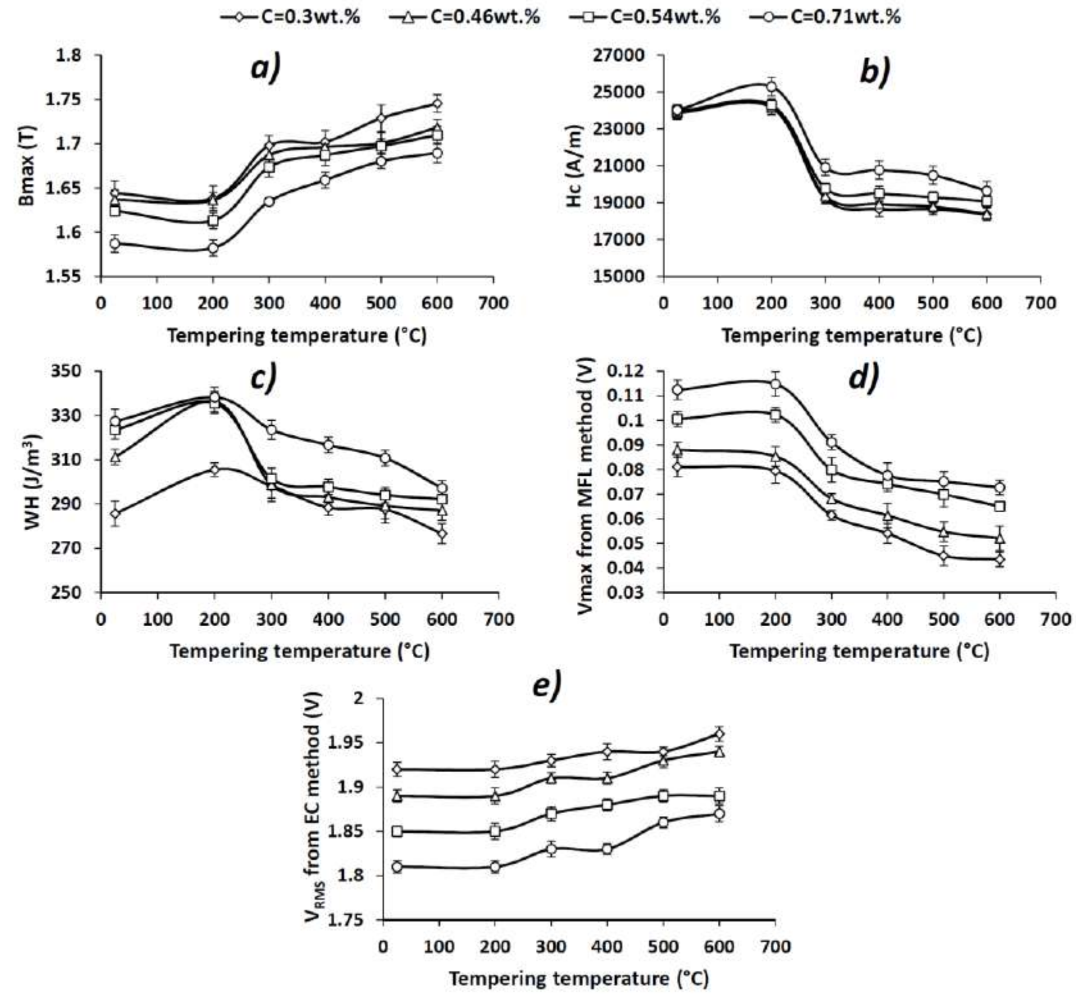


Figure 4. ANN inputs: outputs of NDE tests including a) Bmax from MHL method, b) Hc from MHL method, c) WH from MHL method, d) Vmax from MFL method, and e) voltage RMS from EC method, versus tempering temperature for heat-treated specimens containing different amounts of carbon.

the high dislocations density resulted from shear deformation of martensitic transformation. Therefore, higher magnetic fields are required for the motion of domain walls in the magnetization process. As a result, for the martensite microstructures, MHLs are created in short and wide shapes and as a result, show low values of Bmax and high values of Hc and WH. Besides, due to the direct relationship between Bmax and the amount of flux density near the surface (which is equal to EC output: VRMS) and the inverse relationship between Bmax and flux leaked from the surface (which is equal to MFL output: Vmax), lower values of VRMS and higher amount of Vmax have been obtained for martensitic microstructure. As it can be seen in Figure 4, for all four grades of steels, a slight increase in Hc and WH values have been obtained for the specimens tempered at 200°C. Indeed, the formation of transition carbide (ε-carbides) occurred by tempering at 200°C [25], increases the strength of the pinning sites versus the motion of domain walls in the magnetizing process. This in turn leads to a slight increase in Hc and WH values. At higher tempering temperatures, reduction in residual stresses (crystalline defects especially dislocations density) as well as, the transformation of retained austenite (paramagnetic phase) to the soft ferromagnetic ferrite and bainite (for medium and high carbon grades) facilitate the magnetizing process. Therefore, by tempering from 200 to 600°C, “Bmax and VRMS” values show a continuously increasing trend, while Hc, WH, and Vmax ones exhibit an opposite moderate decrease.

Figure 4 also shows various values of NDE outputs for samples with different carbon content. Indeed, for all the heat-treated groups of samples, an increase in carbon content increases Hc, WH, and Vmax, and reduces Bmax and VRMS. The reverse trend of Bmax (and VRMS) versus carbon content is expected since the increase in carbon content causes an increase in the c/a (lattice tetragonality) ratio which in turn results in a higher cementite fraction after the tempering process. On the other hand, when correlating Hc (and WH) with carbon content in the steels, the tendency was opposite. Indeed, coercivity represents the intensity of the applied magnetic field required to reduce the magnetization of that material to zero. Steel groups containing higher carbon contents showed the higher Hc, because of their weak response to the applied magnetic field due to the high c/a (high distortion due to the confinement of carbon atoms). Therefore, higher carbon contents restrict the domain walls' motion in the magnetization process, and hence, higher applied magnetic fields or in other words, higher coercivities are required to eliminate the magnetization of the steel.

3. 3. Results Obtained from Applying ANN Here in this research, three different magnetic parameters extracted from the MHL method (Bmax, WH, Hc), one parameter from the EC method (VRMS) and another parameter from the MFL method (Vmax) is separately measured. These parameters could be given to a GRNN of Figure 2 to simultaneously determine three mechanical properties of the specimen (including elongation, UTS, and hardness). Each of the five available inputs (Bmax, WH, Hc, VRMS, and Vmax) and all of their combinations, (a total of 31 different cases) may be thought of as the network input. As it can be seen in Figure 5, the accuracy of estimation is not the same for all these combinations; therefore, one should choose an appropriate combination of available inputs for each desired mechanical property. Plots of Figure 5 demonstrates the %RMS error of mechanical properties estimated using the proposed method vs the number of data points used for training for all 31 possible cases. The values not used in the training process are utilized to study the performance of the proposed method. The %RMS error is defined as Equation (1):

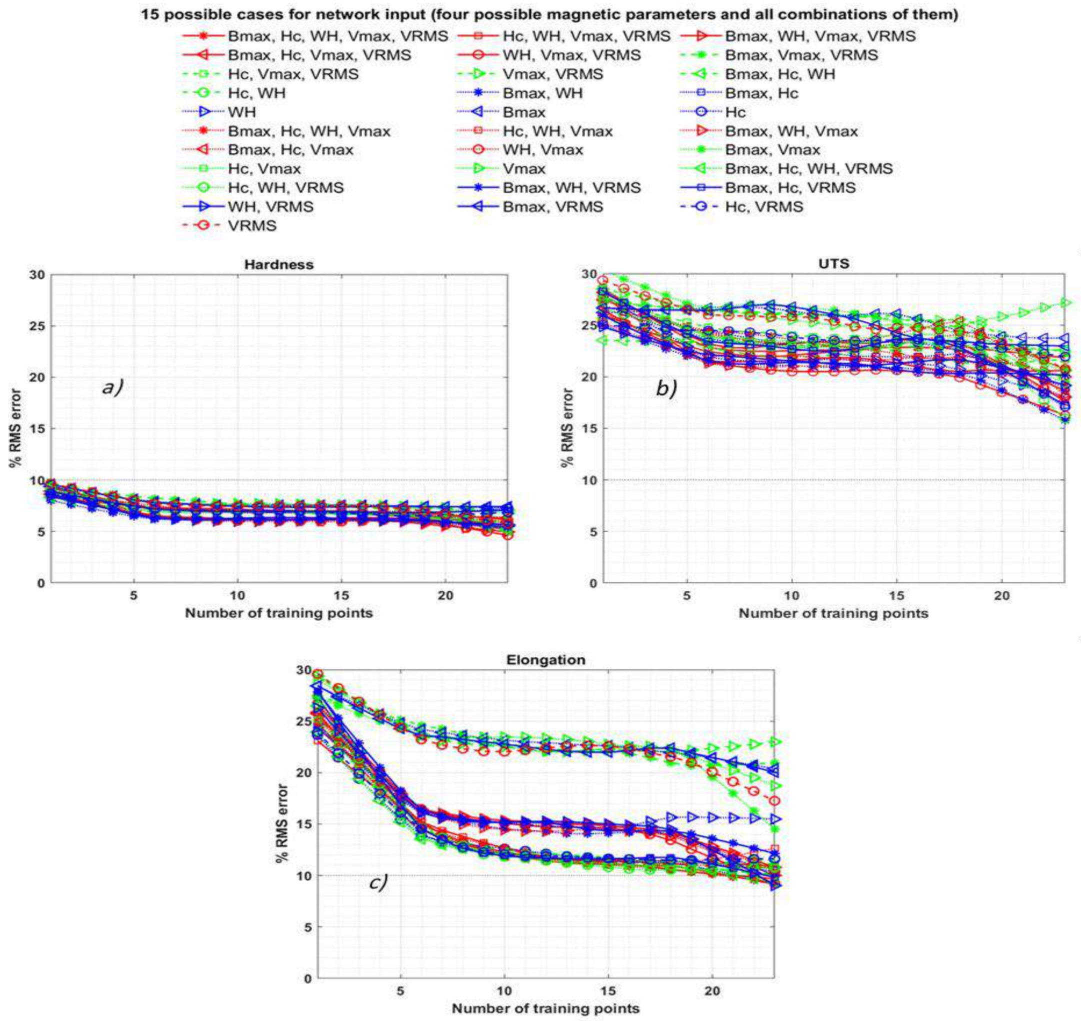


Figure 5. %RMS error of estimated mechanical properties including a) hardness, b) UTS, and c) elongation versus the number of data points used for training the network for all 31 input cases

$$\%RMS\ error\ (N) = 100\frac{\sqrt{\frac{\sum_{n=1}^{(24-N)}(\text{mech. par.}_{n}^{est.} - \text{mech. par.}_{n}^{meas.})^2}{(24-N)}}}{\max(\text{mech. par.}^{meas.})} \quad (1)$$

where N is the size of the dataset used for training the network and (24-N) is the size of the test dataset.

Figure 6 summarizes the accuracy of the method proposed in this work for each of the mechanical parameters estimated and for each input case (from all possible 31 cases). Each category of bars demonstrates the %RMS error of the three estimated mechanical properties (hardness, UTS, elongation) for five available inputs (Bmax, WH, Hc, VRMS, and Vmax) and any combination of them. As it is highlighted by red bars, to estimate elongation, UTS, and hardness, the best input sets are respectively: "Bmax, Hc, WH, and Vmax", "Bmax and WH" and "Bmax, Hc, Vmax and VRMS". Using these magnetic parameters as the network input, the accuracy of hardness, UTS and elongation estimation for 23 training points would be respectively: 4.9, 16.5 and 9.1%. This could also be seen from Figure 5, in which for the hardness, UTS, and elongation, the curves of "Bmax, Hc, Vmax, and VRMS", "Bmax and WH" and "Bmax, Hc, WH, and Vmax" are respectively below the others.

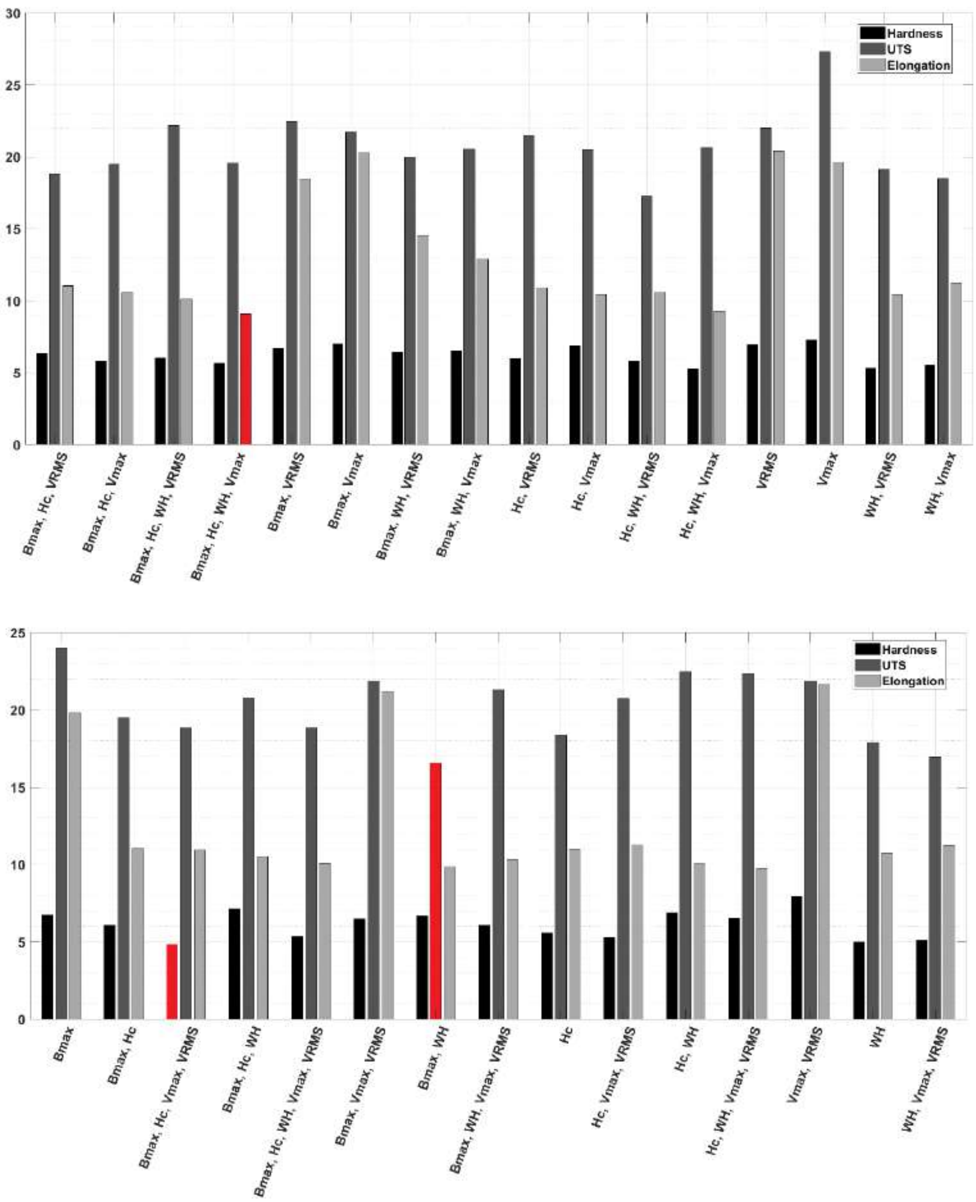


Figure 6. %RMS error (23) for each mechanical property estimated using the proposed method and for each input case

This is an important point and means among all possible cases, one should only use “Bmax, Hc, Vmax and VRMS”, “Bmax and WH” and “Bmax, Hc, WH and Vmax” as the network input.

Although the accuracy of hardness estimation is remarkable and for almost all applications it is completely acceptable for elongation, results do not show a high accuracy for UTS estimation. This could be due to the fact that UTS points (as shown in Figure 2-b) for tempering temperature of 400°C for different carbon contents fall on each other, which in turn reduces the accuracy of prediction. Calibrating the measurement/prediction system with more samples could prevent this problem and improve the results.

4. CONCLUSION

In this work, in the first step, relationships among electromagnetic parameters (Bmax, Hc, WH, VRMS, and Vmax) extracted from MHL, EC, and MFL methods and three important mechanical parameters (hardness, UTS, and elongation) of a set of four heat-treated hypoeutectoid plain carbon steel grades were studied. Based on the results, three proposed NDE techniques including MHL, MFL, and EC methods showed high sensitivity to mechanical properties and carbon content of the plain carbon steels during the tempering process. Reverse trends have been observed for increasing carbon content and tempering temperature versus NDE outputs. In the next step, the mechanical parameters were simultaneously estimated using a generalized neural network (GRNN) with the magnetic parameters as its inputs. The results show that using “Bmax, Hc, Vmax and VRMS”, “Bmax and WH” and “Bmax, Hc, WH, and Vmax” as the network inputs and 23 training data points, 4.9, 16.5, and 9.1% RMS error for estimation of hardness, UTS and elongation.

5. REFERENCES

1. Altpeter, I., R. Tschuncky, and Szielasko, K. Electromagnetic techniques for materials characterization, in the book: Materials Characterization Using Nondestructive Evaluation (NDE) Methods, Elsevier (2016). 225-262. DOI: 10.1016/B978-0-08-100040-3.00008-0
2. Kashefi, M. and Kahrobaee, S. “On the application of non-destructive eddy current method for quality control of heat treated parts”, in 18th Congress IFHTSE, International Federation for Heat Treatment and Surface Engineering, Rio de Janeiro, (2010).
3. Nezhad, K.K., Kahrobaee, S., and Akhlaghi, I. A. “Application of magnetic hysteresis loop method to determine prior austenite grain size in plain carbon steels.” ***Journal of Magnetism and Magnetic Materials,*** Vol. 477, (2019), 275-282. DOI: https://doi.org/10.1016/j.jmmm.2019.01.074
4. Seyfpour, M., Ghanei, S., Mazinani, M., Kashefi, M. and Davis, C. “Nondestructive examination of recovery stage during annealing of a cold-rolled low-carbon steel using eddy current testing technique.” ***Nondestructive Testing and Evaluation***, Vol. 33, No. 2, (2018), 165-174. DOI: 10.1080/10589759.2017.1409746
5. Zhang, C., Bowler, N. and Lo, C. “Magnetic characterization of surface-hardened steel.” ***Journal of Magnetism and Magnetic Materials,*** Vol. 321, No. 23, (2009), 3878-3887. DOI: https://doi.org/10.1016/j.jmmm.2009.07.065
6. Kobayashi, S., Takahashi, H. and Kamada, Y. “Evaluation of case depth in induction-hardened steels: magnetic hysteresis measurements and hardness-depth profiling by differential permeability analysis.” ***Journal of Magnetism and Magnetic Materials,*** Vol. 343, (2013), 112-118. DOI: 10.1016/j.jmmm.2013.04.082
7. Kahrobaee, S., Hejazi, T. H. and Akhlaghi, I. A. “Electromagnetic methods to improve the nondestructive characterization of induction hardened steels: A statistical modeling approach.” ***Surface and Coatings Technology,*** Vol. 380, (2019) 125074. DOI: https://doi.org/10.1016/j.surfcoat.2019.125074
8. Kahrobaee, S., Torbati, M. K. and Amiri, M. S. Surface Characterization of Carburized Steel by Eddy Current, in the book: Nondestructive Testing of Materials and Structures, Springer (2013), 559-564. DOI: 10.1007/978-94-007-0723-8_80
9. Kahrobaee, S., Kashefi, M. and Alam, A. S. “Magnetic NDT Technology for characterization of decarburizing depth.” ***Surface and Coatings Technology,*** Vol. 205, No. 16, (2011), 4083-4088. DOI: https://doi.org/10.1016/j.surfcoat.2011.02.060.
10. Zhu, W., Yin, W., Dewey, S., Hunt, P., Davis, C. L., and Peyton, A. J. “Modeling and experimental study of a multi-frequency electromagnetic sensor system for rail decarburization measurement.” ***NDT & E International,*** Vol. 86, (2017), 1-6. https://doi.org/10.1016/j.ndteint.2016.11.004
11. Falahat, S., Ghanei, S. and Kashefi, M. “Nondestructive examination of decarburised layer of steels using eddy current and magnetic Barkhausen noise testing techniques.” ***Nondestructive Testing and Evaluation,*** Vol. 33, No. 2, (2018), 154-164. DOI: https://doi.org/10.1080/10589759.2017.1397144
12. Kahrobaee, S. and Kashefi, M. “Microstructural characterization of quenched AISI D2 tool steel using magnetic/electromagnetic nondestructive techniques.” ***IEEE Transactions on Magnetics,*** Vol. 51, No. 9, (2015), 1-7. DOI: 10.1109/TMAG.2015.2428673
13. Kahrobaee, S., Ghanei, S. and Kashefi, M. “Using an Artificial Neural Network for Nondestructive Evaluation of the Heat Treating Processes for D2 Tool Steels.” ***Journal of Materials Engineering and Performance,*** Vol. 28, No. 5, (2019), 3001-3011. DOI: 10.1007/s11665-019-04057-4
14. Kahrobaee, S., Norouzi, S. H. and Akhlaghi, I. A. “Nondestructive Characterization of Microstructure and Mechanical Properties of Heat Treated H13 Tool Steel Using Magnetic Hysteresis Loop Methodology.” ***Research in Nondestructive Evaluation,*** Vol. 30, No. 5, (2019), 303-315. DOI: https://doi.org/10.1080/09349847.2019.1574942
15. Mirzaee, A., Kahrobaee, S. and Akhlaghi, I. A. “Non-destructive determination of microstructural/mechanical properties and thickness variations in API X65 steel using magnetic hysteresis loop and artificial neural networks.” ***Nondestructive Testing and Evaluation,*** Vol. 35, No. 72, (2019), 1-17, DOI: 10.1080/10589759.2019.1662901
16. Akhlaghi, I. A., Salkhordeh-Haghighi, M. Kahrobaee, S. and Hojati, M. “Prediction of chemical composition and mechanical properties in powder metallurgical steels using multi-electromagnetic nondestructive methods and a data fusion system.” ***Journal of Magnetism and Magnetic Materials,*** Vol. 498, (2020), 166246. DOI: https://doi.org/10.1016/j.jmmm.2019.166246

17. Kahrobaee, S. and Zohourvahid-Karimi, E. "Characterisation of work-hardening in Hadfield steel using non-destructive eddy current method." ***Nondestructive Testing and Evaluation,*** Vol. 34, No. 2, (2019), 178-192. DOI: https://doi.org/10.1080/10589759.2019.1581190
18. Tandon, K., "MFL tool hardware for pipeline inspection." ***Materials Performance,*** Vol. 36, No. 2, (1997), 75-79. DOI: 10.1016/s0963-8695(98)90832-5
19. Haines, H., Porter, P. C., and Barkdull, L. "Advanced MFL signal analysis aids pipe corrosion detection." ***Pipe Line & Gas Industry,*** Vol. 82, No. 3, (1999), 49-57.
20. Li, Y., Tian, G. Y. and Ward, S. "Numerical simulation on magnetic flux leakage evaluation at high speed." ***NDT & E International,*** Vol. 39, No. 5, (2006) 367-373. DOI: https://doi.org/10.1016/j.ndteint.2005.10.006
21. Ricken, W., Schoenekess, H. and Becker, W. J. "Improved multi-sensor for force measurement of pre-stressed steel cables by means of the eddy current technique." ***Sensors and Actuators A: Physical,*** Vol. 129, No. 1-2, (2006), 80-85. DOI: https://doi.org/10.1016/j.sna.2005.11.056
22. Hassoun, M. H., Fundamentals of artificial neural networks. MIT Press, 1995.
23. Aggarwal, C. C., Neural networks and deep learning. Springer, 2018.
24. Specht, D. F., "A general regression neural network." ***IEEE Transactions on Neural Networks,*** Vol. 2, No. 6, (1991), 568-576. DOI: 10.1109/72.97934
25. Handbook, A., Heat treating. ASM International Materials Park (OH), 1991.

Persian Abstract

چکیده

در این مقاله، نتایج حاصل از اعمال یک شبکه عصبی مصنوعی بر خروجی‌های حاصل از سه آزمون غیرمخرب مغناطیسی شامل حلقه هیسترزیس مغناطیسی، جریان گردابی و نشت شار مغناطیسی، به منظور تعیین خواص مکانیکی فولادهای ساده کربنی بازپخت شده با درصد کربن نامعلوم ارائه شده است. به منظور بررسی همزمان تاثیر درصد کربن و ریزساختار بر خواص مغناطیسی و مکانیکی قطعات فولادی، چهار گروه از فولادهای هیپویوتکتوید حاوی 0/30، 0/45، 0/54 و 0/71 درصد وزنی کربن در دمای °C 950 آستنیته و سپس تحت عملیات کوئنچ و بازپخت در محدوده دمایی 200-600°C قرار گرفتند. پس از عملیات حرارتی، توسط آزمون کشش و سختی‌سنجی خواص مکانیکی قطعات شامل استحکام کششی، انعطاف‌پذیری و سختی اندازه‌گیری شد. در نهایت به منظور بررسی خواص الکترومغناطیسی قطعات، آزمون‌های غیرمخرب حلقه هیسترزیس مغناطیسی، جریان گردابی و نشت شار مغناطیسی انجام و از خروجی این آزمون‌ها به عنوان ورودی شبکه عصبی طراحی شده استفاده شده است. نتایج حاصل نشان می‌دهد که در صورت استفاده از ترکیب مناسبی از خواص الکترومغناطیسی به عنوان ورودی شبکه عصبی، امکان تعیین سختی، استحکام کششی و درصد ازدیاد طول فولادهای هیپویوتکتوئید پس از عملیات بازپخت با دقت بالا وجود دارد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Experimental Investigation of Uncoated Electrode and PVD AlCrNi Coating on Surface Roughness in Electrical Discharge Machining of Ti-6Al-4V

N. H. Phan*[a], P. V. Dong[a], T. Muthuramalingam[b], N. V. Thien[a], H. T. Dung[a], T. Q. Hung[a], N. V. Duc[a], N. T. Ly[a]

[a] *Hanoi University of Industry, Cau Dien Street, Bac Tu Liem District, Hanoi, Vietnam*
[b] *Department of Mechatronics Engineering, SRM Institute of Science and Technology, Kattankulathur, India*

PAPER INFO

Paper history:
Received 21 December 2020
Received in revised form 24 January 2021
Accepted 31 January 2020

Keywords:
EDM
Regression Model
Taguchi
Coated Electrode

ABSTRACT

The surface texture on the EDM is an important quality indicator since it directly affects the cost of the further finishing work. The coating over the tool electrode in EDM can improve productivity, electrode wear resistance and surface quality. In the present study, the surface roughness of the EDM machined surface with coated and uncoated electrodes was evaluated. Al and AlCrNi coated Al electrode has been used for the study on machining Titanium alloy (Ti-6Al-4V). Current (I), voltage (V_g) and pulse on time (Ton) have been used as technology parameters under Taguchi method with regression model and optimal technology parameters. It was found as I and V_g are the parameters could strongly affect the surface quality. The coated tool electrode can produce better surface quality than uncoated tool electrode. The optimal technological parameters with coated and uncoated electrodes were found as I = 10 A, T_{on} = 500 µs and V_g = 40 V.

doi: *10.5829/ije.2021.34.04a.19*

1. INTRODUCTION

Titanium and its alloys are widely used in various fields including aerospace, nuclear, medical, etc. It is highly difficult to process such materials with traditional machining methods due to the higher strength. Electrical discharge machining (EDM) is commonly used to process this group of materials. In EDM, the machining surface quality and the machining productivity is not high [1]. Many technical solutions have been implemented to improve the machining efficiency of EDM including optimization of technological parameters, new electrode material or electrode surface layer and vibration in EDM. It has shown that the utilization of coated electrodes EDM is a great promising solution to improve machining productivity and quality. It can reduce the wear of tool electrode and production costs. However, research works on EDM using new electrode materials are mainly focused with powder metallurgical electrodes. The usage of coating electrodes are still very little. In addition, the physical and mechanical properties of the material layer of the electrode surface will directly affect the spark formation process to enhance the machining productivity and quality. Hence it is necessary to clarify the effectiveness of using coated electrodes in EDM process.

Extensive research were performed to evaluate the machinability of the electrode in EDM [2]. Many types of electrode materials (Al, Cu, Cu-W and brass) have been investigated in tool steel machining by EDM [3]. The Cu-W could produce lower surface roughness (R_a) in EDM [4]. The machining capacity with Cu-W electrode is also higher than Cu and brasselectrodes [5]. It was inferred that the formation of larger residual tensile stress in EDM process with conventional electrodes [6]. The electrode material used in EDM would affect the structure of the white layer on the machined surface layer consisting of austenite and residual stress [7, 8]. The different electrode materials are used in EDM to modify the surface quality [9 -11]. The R_a and morphology of the machining surface are strongly influenced by the changes of electrode material in EDM [12]. Recent studies have used powder metallurgical electrodes in EDM to improve

*Corresponding Author Institutional Email: phanktcn@gmail.com (N. H. Phan)

Please cite this article as: N. H. Phan, P. V. Dong, T. Mthuramalingam, N. V. Thien, H. T. Dung, T. Q. Hung, N. V. Duc, N. T. Ly, Experimental Investigation of Uncoated Electrode and PVD AlCrNi Coating on Surface Roughness in Electrical Discharge Machining of Ti-6Al-4V, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 928-934

machining efficiency with improved surface quality [13]. The TiC composite electrode can produce better surface than Cu-W electrode [14]. The powder metallurgical electrode can produce alloy layer with better surface quality and hardness[15].The composite electrodes in EDM can be fabricated with 3D printing [16]. However, it should be required that the composition of the composite mixture of the electrode material should be corrected since it would directly affect the surface quality and machinability in EDM [17].

The surface coating technology is very popular to create a different surface material from the base material, and this improves the workability of the products. Several types of electrode materials (Cu, brass, molyden and Cu-W) coated with nickel and diamond - nickel materials have been used to investigate the improvement of TWR in EDM [18]. TWR is greatly reducedin EDM with coated electrode such as diamond - nickel coated electrode. I and T_{on} are significantly affect MRR and EWR in EDM with coated electrode [19]. The regression model is determined by Taguchi - GRA, it has consistent accuracy.The higher wear resistance of the in EDM was observed with zinc coated electrode [20]. As compared to Cu electrodes, TiN or TiAlN coated Cu electrodes would be suitable for EDM better finishing [21]. The influence of parameters including V_g, I and T_{on} on the quality indicators in EDM was investigated. It was found that I was the most significant influencial parameter. The coated electrode can significantly improve MRR, TWR, R_a and machining accuracy [22]. The roughness of the machined surface with the Cu electrode has been selected to evaluate the effectiveness of coated Cu electrode including silver, nickel, zinc and epoxy. It was found that the coated electrode can significantly improve the surface quality [23]. The zinc coated electrode provides more efficient solution than many other methods [24]. The optimal values of MRR, TWR and R_a in EDM using silver coated electrode were determined by Taguchi [25]. The optimization and regression model analysis was found with accepted accuracy of practice [26]. Taguchi method is commonly used to design experiments and analyze the effect of technology parameters on quality in the EDM using powder metallurgy electrode and coated electrode [27]. Taguchi method can be very suitable for research on the effect of new coating materials on the electrode surface on the quality parameters in EDM.

The surface roughness in EDM is an indicator directly related to the choice of the further finishing method and its machining cost. Hence a research attempt is very necessary to improve surface quality in EDM. The number of studies regarding coating electrodes and the coating materials used are still very limited. In this study, the effect of technological parameters on surface roughness (R_a) in EDM using uncoated and AlCrNi coated Al electrode was investigated. The regression model development in Taguchi has identified the regression equation of the R_a. The efficiency of the coating electrode to improve the R_a and the quality of the machining surface at the optimum condition were also analyzed.

2. EXPERIMENTAL METHODOLOGY

The machining experiments were performed on CNC type Electro Discharge Machine manufactured by Electronics India Private Limitedfor machining Ti-6Al-4V alloy as shown in Figures 1 and 2. The process variables were chosen based on low, medium and high level of process parameter available at machine. All experiments were systematically planned with four level based on Taguchi method. The levels of process parameters in this study described in Table 1. Since the work deals with three factors and four levels, L16 orthogonal array (OA) was selected in the present study to evaluate quality measurement of surface roughness (R_a). Ra of machined workpiece surface was measured by contact type surface roughness tester (Taylor Hobson machine) with the cut off length of 0.8mm.

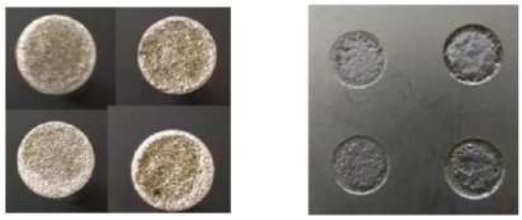
3. RESULTS AND DISCUSSION

3. 1. Effects of Process Parameters on R_a using Analysis of Variance (ANOVA)

The ANOVA analysis of R_a results to find influence of parameters on quality indicators in EDM is shown inTable 2. I (F = 161.13 with Al electrode and F = 177.69 with AlCrNi coated electrode) and V_g (F = 6.19 with Al electrode and F = 8.25 with AlCrNi coated electrode) had a significant effect on R_a in EDM with Al electrode and AlCrNi coated

(a) Electrode (b) Workpiece

Figure 1. Machined surface using EDMwith Al electrode



(a) Electrode (b) Workpiece

Figure 2. Machined surface using EDM with AlCrNi electrode

TABLE 1. Experimental results from the present study

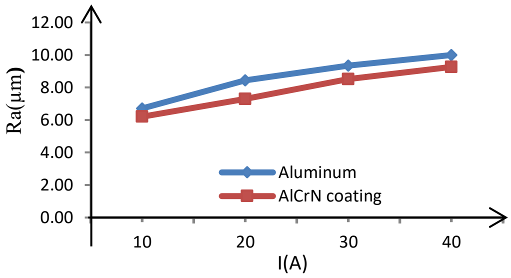
Input process parameters			R_a (µm)				Relative error (%)	
			Experimental		Calculated			
I (A)	T_{on} (µs)	V_g (V)	Al	AlCrNi coating	Al	AlCrNi coating	Al	AlCrNi coating
10	100	40	6.664	6.111	6.715	5.954	0.765	-2.576
10	500	45	6.683	6.162	6.914	6.159	3.452	-0.043
10	1000	50	6.691	6.251	7.108	6.361	6.234	1.764
10	1500	55	6.815	6.294	7.303	6.563	7.155	4.277
20	500	40	7.781	6.874	7.774	6.979	-0.085	1.533
20	100	45	8.325	6.976	8.008	7.217	-3.814	3.448
20	1500	50	8.662	7.648	8.163	7.383	-5.758	-3.463
20	1000	55	8.981	7.669	8.401	7.624	-6.461	-0.583
30	1000	40	9.116	8.045	8.829	8.001	-3.143	-0.544
30	1500	45	9.203	8.338	9.024	8.203	-1.946	-1.617
30	100	50	9.412	8.768	9.300	8.480	-1.189	-3.290
30	500	55	9.665	8.946	9.499	8.685	-1.720	-2.914
40	1500	40	9.706	9.013	9.885	9.023	1.840	0.112
40	1000	45	9.783	9.112	10.122	9.264	3.465	1.671
40	500	50	10.112	9.313	10.359	9.505	2.447	2.065
40	100	55	10.391	9.623	10.593	9.743	1.940	1.242

TABLE 2. ANOVA of R_a using Aluminum electrode and AlCrNi coated electrode

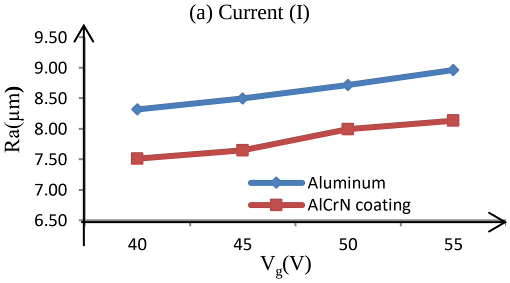
Source	DF	SS		MS		F-Value		P-Value		Contribution%	
		Al	AlCrNi	Al	AlCrNi	Al	AlCrNi	Al	AlCrNi	Al	AlCrNi
I	3	24.3974	21.8944	8.13246	7.29813	161.13	177.69	0	0	95	94.5
Vg	3	0.9366	1.0166	0.3122	0.33887	6.19	8.25	0.029	0.015	3.64	4.4
Ton	3	0.0426	0.0202	0.0142	0.00672	0.28	0.16	0.837	0.917	0.16	0.09
Error	6	0.3028	0.2464	0.05047	0.04107	-	-	-	-	0.01	1.01
Total	15	25.6794	23.1776				-				

electrode, and T_{on} (F = 0.28 with Al electrode and F = 0.16 with AlCrNi coated electrode) had a negligible effect on R_a. Based on the values of the fisher coefficient (F), I is the most significant effect (F is largest), followed by V_g and T_{on} (F is minimum), respectively. The percentage distribution of the influence of the technological parameters on R_a for both types of electrode materials is quite similar. The percentage of distribution of the influence of I is the largest (95% withAl electrode and 94.5% with AlCrNi coated electrode), it's of V_g (3.65% with Al electrode and 4.4.5% with AlCrNi coated electrode) and T_{on} is very small (0.16% with Al electrode and 0.09% with AlCrNi coated electrode).

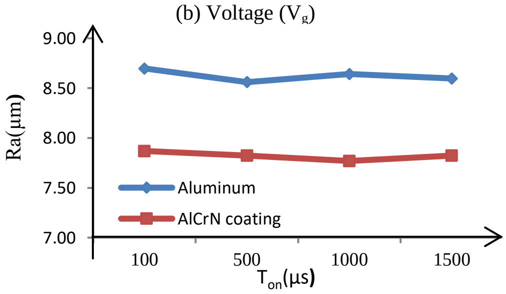
The variation of current (I), voltage (V_g) and pulse on time (T_{on}) has affected R_a in EDM with Al electrode and AlCrNi coated electrode is shown in Figure 3. The higher I and V_g has lead to increase in R_a as shown in Figures 3a and 3b. The R_a was increased due to the increase of I and V_g which has led to the higher heating energy to produce larger size and depth of the craters on the machining surface [25]. The size and number of adhesion particles have been increased with larger melting and evaporation of the electrode material and workpiece [6]. As compared to R_a at I = 10 A, the maximum increase in R_a at I = 40A was 48.9% with the Al electrode and 49.3% with AlCrNi coated electrode. The 8.3% lower R_a has been found with AlCrNi coated electrode. The influence of I on R_a is much greater than V_g.The changes in T_{on} has resulted in negligible change in R_a with AlCrNi coated electrode as shown in Figure 3c. The R_a in EDM with Al electrode was minimum at T_{on} = 500 µs. The R_a was minimum at



(a) Current (I)



(b) Voltage (V_g)



(c) Pulse on time (T_{on})

Figure 3. The influence of technology parameters on R_a

T_{on} = 1000 µs with AlCrNi coated electrode. The higher T_{on} led to an increase in the spark discharge channel. However the lower peak of the spark would lead to a decrease in R_a [7].

Figure 3 also shows that R_a in EDM using AlCrNi coated electrode was smaller than with Al electrode. This shows that the machining surface quality has been improved with the coating electrode. The change of I leads to the greatest difference between the R_a of the two electrodes. The reason for the lower R_a in EDM using AlCrNi coated electrode was due to the higher thermal stability of AlCrNi coating material on the surface. This has led to higher erosive resistance of the electrode surface by the heat of the sparks. Hence the surface layer of the electrode coated with AlCrNi could be changed. The surface hardness of the electrode could also affect the roughness of the machining surface [27]. Since the coating affects the electrical conductivity of spark plasma column, it affects the discharge energy delivered per every pulse. It has considerably affected the surface roughness [27]. The difference between the thermal and electrical conductivity characteristics of the coated electrode material could affect the process of the spark formation and the energy of the generated sparks [22].

3. 2. Effects of Process Parameters on R_a using Linear Regression Model

Many methods have been used to establish regression models in EDM including Taguchi, RSM, ANN, etc. It has been inferred that EDM process requires statistical analysis and optimization to obtain optimal process parameters combinations during the machining process for better surface quality [28]. In this study, Taguchi method was used to determine the regression model of R_a in EDM using Al electrode and AlCrNi coated electrode. The analytical results on the accuracy of the regression model of R_a are shown in Tables 3 and 4. The coefficients R^2 and R^2 (adj) showed the appropriateness of the regression model. Equations (1) and (2) represent a regression model of R_a with Al electrode and AlCrNi coated electrode. The result of the R_a determined by calculation by the regression model is compared with it experimentally (Table 1 and Figure 4). It showed that the experimental and calculated maximum error of R_a was 7.15% with Al electrode as that of 4.2% with AlCrNi coated electrode. The regression model has the accuracy consistent with the experiment. The modeling has also shown that I and V_g were the main influencinal parameters on R_a in EDM using Al electrode and AlCrNi coated electrode.

$$\text{Aluminum} = 3.91504 + 0.10766*I - 4.3465e\text{-}005*T_{on} + 0.04319*V_g \tag{1}$$

$$\text{AlCrNi coating} = 3.14 + 0.104*I - 0.000039*T_{on} + 0.0443*V_g \tag{2}$$

3. 3. Determination of the Optimal Technology Parameters

The S/N coefficient is used to determine the optimal technology parameters. The S/N of the R_a is determined by "smaller is better" as shown in Figure 5. It has been shown that the optimum

TABLE 3. Results of the ANOVA assessment for the linear regression modelof Al electrode

Source	DF	SS	MS	F	P
Regression	3	24.1224	8.0408	61.97	0.000
Residual Error	12	1.5570	0.1297	-	-
Total	15	25.6794	-	-	-
R^2 = 93.9% R^2(adj) = 92.4%					

TABLE 4. Results of the ANOVA assessment for the linear regression model of AlCrNi coated electrode

Source	DF	SS	MS	F	P
Regression	3	22.6813	7.5604	182.79	0.000
Residual Error	12	0.4963	0.0414	-	-
Total	15	23.1776	-	-	-
R^2 = 97.9% R^2(adj) = 97.3%					

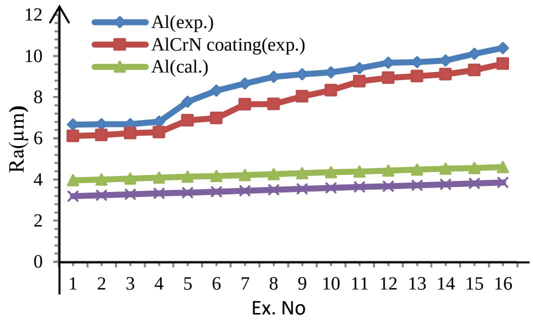
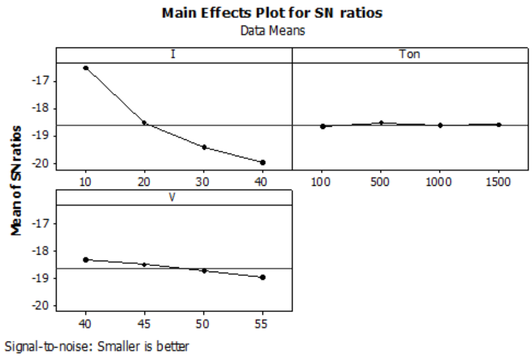
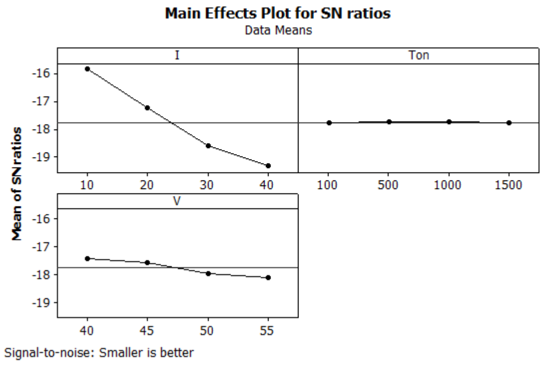


Figure 4. Compare results of experiment and calculated by regression model



(a) Al electrode



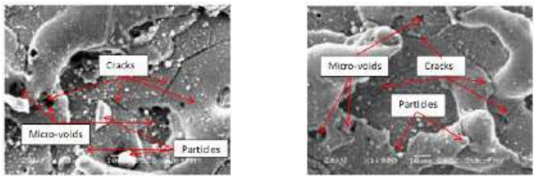
(b) AlCrNi coated electrode

Figure 5. Main effects plot for S / N ratio of R_a

technological parameters in EDM using Al electrode and AlCrNi coated electrode are the identical as o I = 10 A, T_{on} = 500 µs and V_g = 40 V. Based on the regression model Equations (1) and (2), the optimal values for both R_a could be determined as follows $R_{a:Al}$ = 6.698 µm and $R_{a:AlCrNicoated}$ = 5.938 µm.A confirmation experiment at optimum conditions was performed and found as $R_{a:Al}$ = 6.476 µm and $R_{a:AlCrNicoated}$ = 5.858 µm. It was inferred that the error between the calculated method and experimental error was 7.2% with Al electrode and 4.3% with AlCrNi coated electrode.

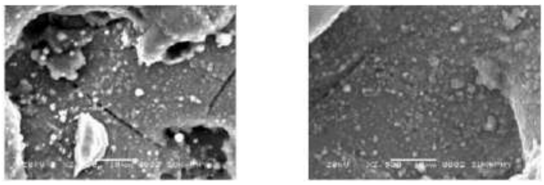
3. 3. Surface Quality Analysis at Optimal Conditions

Figure 6 shows that the lower adhesion debris and micro- cracks on the machined surface in EDM has found with coated electrode as that of Al electrode. The size of particles and the micro - voids on the surface machined with coated electrode was smaller. It could also resulted in the uniform surface topography with lower surface roughness of machined surface with the AlCrNi coated electrode than Al electrode as shown in Figure 7. The size of microscopic cracks on the machined surface with Al electrode was large as shown in Figure 8. This may be due to the improved surface hardness and heat resistance properties of the coating material. It has also resulted in an increase in the erosion resistance of the AlCrNi coated electrode's surface layer with uniformly distributed surface sparks. The electrical and thermal conductivity characteristics of the coating material can facilitate the formation of sparks [24]. Hence the sparks are more uniformly distributed with the smaller beam energy. A partial arcing could be occurred with coated electrode to create tiny micro-voids.



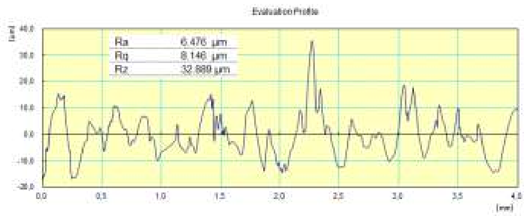
(a) Al electrode (b) AlCrNi coated electrode

Figure 6. SEM micro-graphs of machined surface in EDM

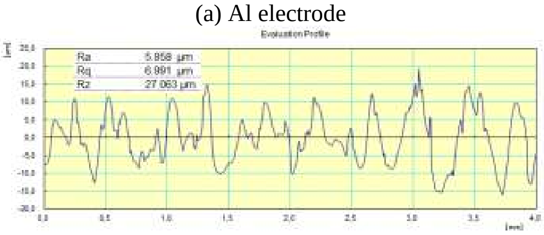


(a) Al electrode (b) AlCrNi coated electrode

Figure 7. Topography of machined surface in EDM



(a) Al electrode



(b) AlCrNi coated electrode

Figure 8. Profile of machined surface in EDM

The discharge gap size was also affected by the surface layer material of the electrode. Hence the size and number of particles adhering to the machining surface could also be improved as the discharge gap size increases. The surface quality after EDM process using AlCrNi coated electrode was observed better than Al electrode.

4. CONCLUSION

In the present study, an investigation was attempted to evaluate R_a of the EDM machined surface with coated and AlCrNi coated Al electrode. Ti-6Al-4V was utilized as specimens using regression model to get optimal technology parameters. From the detailed investigation, the following conclusions were drawn

- The AlCrNi coated tool electrode can produce better surface quality with lower surface roughness, micro cracks and voids than uncoated tool electrode due to its electrical conductivity and melting point.
- The optimal technological parameters with coated and uncoated electrodes were found as I = 10 A, T_{on} = 500 µs and V_g = 40 V with good accuracy of 4.3%.
- I and V_g are the parameters could strongly affect surface quality.
- However more clear research results are needed to comprehensively evaluate the economic and technical efficiency between coated and uncoated electrodes in EDM process.

5. REFERENCES

1. Qosim. N., Supriadi. S., Puspitasari. P., Kreshanti. P., "Mechanical Surface Treatments of Ti-6Al-4V Miniplate Implant Manufactured by Electrical Discharge Machining", ***International Journal of Engineering, Transactions A: Basics,*** Vol. 31, No. 7, (2018), 1103-1108. DOI: 10.5829/ije.2018.31.07a.14
2. Amorim, Fred L., and Walter L. Weingaertner "The Behavior of Graphite and Copper Electrodes on the Finish Die-Sinking Electrical Discharge Machining (EDM) of AISI P20 Tool Steel", ***Journal of the Brazilian Society of Mechanical Sciences and Engineering***, Vol. 29, No. 4, (2007), 366-371. http://dx.doi.org/10.1590/S1678-58782007000400004
3. Naveen. B., Maheshwari. S., Sharma. C., Anil. K., "Performance Evaluation of Powder Metallurgy Electrode in Electrical Discharge Machining of AISI D2 Steel Using Taguchi Method", ***International Journal of Mechanical and Mechatronics Engineering***, Vol. 2, No. 2, (2008), 225-229. DOI: doi.org/10.5281/zenodo.1327508
4. Rahul M., Hare K., Ankit K. S., Ranjan K. G., "A Review on Copper and its alloys used as electrode in EDM", IOP Conf. Series: Materials Science and Engineering, Vol. 377, (2018), 012183. doi:10.1088/1757-899X/377/1/012183.
5. Qudeiri, J. E. A., Zaiout, A., Mourad, A. H. I., Abidi, M. H., & Elkaseer, A. "Principles and Characteristics of Different EDM Processes in Machining Tool and Die Steels", ***Applied Science Journal***, Vol. 10, (2020), 2082. DOI:10.3390/app10062082.
6. Dewangan, S., Biswas, C. K., & Gangopadhyay, S., "Influence of Different Tool Electrode Materials on EDMed Surface Integrity of AISI P20 Tool Steel", ***Materials and Manufacturing Processes***, Vol. 29, No. 11-12, (2014), 1387-1394. DOI: 10.1080/10426914.2014.930892.
7. Ranjith. R, Tamilselvam. P, Prakash. T, Chinnasamy. C, "Examinations concerning the electric discharge machining of AZ91/5B4CP composites utilizing distinctive electrode materials", ***Materials and Manufacturing Processes,*** Vol. 34, No. 10, (2019), 1120-1128. DOI: 10.1080/10426914.2019.1628258.
8. Moudood, M. A., Sabur, A., Lutfi, A., Ali, M. Y., Jaafar, I. H., "Investigation of the machinability of non-conductive ZrO2 with different tool electrodes in EDM", ***International Journal of Automotive and Mechanical Engineering***, Vol. 10, (2014), 1866-1876. DOI: 10.15282/ijame.10.2014.4.0155
9. Patel, S. S., Prajapati, J. M., "Experimental Investigation of Surface Roughness and Kerf Width During Machining of Blanking Die Material on Wire Electric Discharge Machine", ***International Journal of Engineering, Transactions A: Basics*** Vol. 31, No. 10, (2018), 1760-1766. DOI: 10.5829/ije.2018.31.10a.19
10. Pawar, R. S., Sharma, S. B., "Development of EDM Electrode by Powder Metallurgy Process", ***Trends in Mechanical Engineering & Technology***, Vol. 3, No. 2, (2013), 10-15.
11. Saebnoori, E., Vali, I., "Surface Activation of NiTi Alloy By Using Electrochemical Process For Biomimetic Deposition Of Hydroxyapatite Coating", ***International Journal of Engineering, Transactions A: Basics*** Vol. 27, No. 10, (2014), 1627-1634. DOI: 10.5829/idosi.ije.2014.27.10a.17
12. Das, A., Misra, J. P., "Experimental investigation on surface modification of aluminum by electric discharge coating process using TiC/Cu green compact tool-electrode", ***Machining Science and Technology: An International Journal***, Vol. 16, No. 4, (2012), 601-623. https://doi.org/10.1080/10910344.2012.731951
13. Prabharan. C., Thangadurai. K., Gobal. P., Vadivel. S, "Design and Development of EDM Process Electrode through Powder Metallurgy Route", ***International Journal of Engineering Research and Technology***, (2019). ID: 182234588
14. Ndaliman, M. B., Khan, A. A., "Development of Powder Metallurgy (PM) Compacted Cu-TaC Electrodes for EDM", ***Journal of Mechanics Engineering and Automation***, Vol. 1, (2011), 385-391.
15. Liu, Y., Wang, W., Zhang, W., Ma, F., Yang, D., Sha, Z., & Zhang, S., "Experimental Study on ElectrodeWear of Diamond-NickelCoated Electrode in EDM Small Hole Machining", ***Advances in Materials Science and Engineering***, Vol. 2019, (2019), 1-10. https://doi.org/10.1155/2019/7181237
16. Hutsaylyuk, V., Lytvynenko, I., Maruschak, P., Dzyura, V., Schnell, G., Seitz, H., "A New Method for Modeling the Cyclic Structure of the Surface Microrelief of Titanium Alloy Ti6Al4V After Processing with Femtosecond Pulses". ***Materials,*** Vol. 13, (2020), https://doi.org/10.3390/ma13214983.
17. Panchal, D. L., Biradar, S. K., Gosavi, V. Y., "Analysis of EDM Process Parameters by Using Coated Electrodes", ***International Journal of Engineering Trends and Technology***, Vol. 41, No. 4, (2016), 181-185. DOI: 10.14445/22315381/IJETT-V41P234
18. Yuangang, W., Fuling, Z., Jin, W., "Wear-resist Electrodes for Micro-EDM", ***Chinese Journal of Aeronautics***, Vol. 22, (2009), 339-342. https://doi.org/10.1016/S1000-9361(08)60108-9
19. Jothimurug, R., Amirthagad, K. S., Daniel, J., "Performance of Silver Coated Copper Tool with Kerosene-servotherm Dielectric in EDM of Monel 400TM". ***Journal of Applied Sciences***, Vol. 12, (2012), 999-1005. DOI: 10.3923/jas.2012.999.1005
20. Jin, Z. J., Zhang, M., Guo, D. M., Kang, R. K., "Electroforming of Copper/ZrB2 Composite Coating and its Performance as Electro-discharge Machining Electrodes", ***Key Engineering***

Materials, Vol. 291-292, (2005), 537-542. https://doi.org/10.4028/www.scientific.net/KEM.291-292.537

21. Anthuvan, R. N., Krishnaraj, V., "Efect of coated and treated electrodes on Micro‑EDM characteristics of Ti‑6Al‑4V", ***Journal of the Brazilian Society of Mechanical Sciences and Engineering***, Vol. 42, No. 517, (2020). https://doi.org/10.1007/s40430-020-02578-x.
22. Siva RAO, S. V. V. N., Bhavani, T., Nath, R. K., Maji, P., Ghosh, S. K., Barma, J. D., "Surface modification by electro discharge machining using powder metallurgy electrode: A review", ***Surface Review and Letters***, Vol. 28, No. 01, 2030004 (2021). doi: 10.1142/S0218625X2030004X
23. Rashid, A., Bilal, A., Liu, C., Jahan, M. P., Talamona, D., Perveen, A., "Effect of Conductive Coatings on Micro-Electro-Discharge Machinability of Aluminum Nitride Ceramic Using On-Machine-Fabricated Microelectrodes", ***Materials,*** Vol. 12, 3316, (2019). DOI:10.3390/ma12203316
24. Chanin, M. N., Kuei, C. H., Lin, C., "Using Taguchi design, regression analysis and simulation to study maintenance float systems", ***International Journal of Production Research***, Vol. 28, No. 11, (1990), 1939-1953. DOI:10.1080/00207549008942844
25. Patowari, P. K., Saha, P., Mishra, P. K., "Taguchi analysis of surface modification technique using W-Cu powder metallurgy sintered tools in EDM and characterization of the deposited layer", ***The International Journal of Advanced Manufacturing Technology,*** Vol. 54, (2011), 593–604. DOI: 10.1007/s00170-010-2966-y.
26. Marafona, J. D., Araujo, A., "Influence of workpiece hardness on EDM performance", ***International Journal of Machine Tools & Manufacture***, Vol. 49, (2009), 744-748.
27. Muthuramalingam, T., Mohan, B., "Performance analysis of iso current pulse generator on machining characteristics in EDM process", ***Archives of Civil and Mechanical Engineering***, Vol. 14, No. 3, (2014), 383-390.
28. Muthuramalingam, T., Saravanakumar, D., Babu, L. G., Phan, N. H., Pi, V. N., "Experimental investigation of white layer thickness on EDM processed silicon steel using ANFIS approach", ***Silicon,*** Vol. 12, No. 8, (2020), 1905-1911.

Persian Abstract

چکیده

بافت سطح بر روی EDM یک شاخص کیفیت مهم است زیرا مستقیماً بر هزینه کار تمام شده بعدی تأثیر می گذارد. پوشش روی الکترود ابزار در EDM می تواند باعث بهبود بهره وری ، مقاومت در برابر سایش الکترود و کیفیت سطح شود. در مطالعه حاضر ، زبری سطح ماشینکاری شده EDM با الکترودهای روکش دار و بدون روکش بررسی شد. برای مطالعه در ماشینکاری آلیاژ تیتانیوم (Ti-6Al-4V)از الکترود Al با روکش Al و AlCrNi استفاده شده است. جریان (I)، ولتاژ (Vg)و پالس به موقع (تن) به عنوان پارامترهای فناوری تحت روش تاگوچی با مدل رگرسیون و پارامترهای بهینه فناوری استفاده شده اند. از آنجا که I و Vg پارامترهایی هستند که می توانند به شدت بر کیفیت سطح تأثیر بگذارند ، مشخص شد. الکترود ابزار روکش دار می تواند کیفیت سطح بهتری نسبت به الکترود ابزار بدون روکش تولید کند. پارامترهای فن آوری بهینه با الکترودهای پوشش داده شده و بدون پوشش به عنوان I = 10 A، Ton = 500 میکرو ثانیه و Vg = 40 V پیدا شد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 935-947

International Journal of Engineering

Journal Homepage: www.ije.ir

Robust and Stable Flow Shop Scheduling Problem under Uncertain Processing Times and Machines' Disruption

Z. Abtahi , R. Sahraeian*

Department of Industrial Engineering, College of Engineering, Shahed University, Persian Gulf Expressway, Tehran, Iran

P A P E R I N F O

Paper history:
Received 31 August 2020
Received in revised form 20 November 2020
Accepted 03 December 2020

Keywords:
Machine Breakdowns
Processing Time Variability
Robust
Scheduling
Stable
Two Machine Flow Shop System

A B S T R A C T

This paper presents a predictive robust and stable approach for a two-machine flow shop scheduling problem with machine disruption and uncertain job processing time. Indeed, a general approach is proposed that can be used for robustness and stability optimization in an m-machine flow shop or job shop scheduling problem. The robustness measure is the total expected realized completion time. The expected sum of squared aberration between each jobs' completion time in the realized and initial schedules is the stability measure. We proposed and compared two methods to deal with such an NP-hard problem; a method based on decomposing the problem into sub-problem and solving each sub-problem, and a theorem-based method. The extensive computational results indicated that the second method has a better performance in terms of robustness and stability, especially in large-sized problems. In other words, the second method is preferable because of the better manufacturer responsiveness to the customer and the production staff satisfaction enhancement.

doi: *10.5829/ije.2021.34.04a.20*

NOMENCLATURE

D	Downtimes (a General distribution $D \sim G(t)$); the time required to back the machine to the operational mood	λ_j	The exponential distribution rate of generation initial processing time of job *j* on machine *1*
U	Uptimes (an exponential distribution with rate θ); The time between two consecutive machine breakdowns	μ_j	The exponential distribution rate of generation initial processing time of job *j* on machine *2*
i	Machine index, $i = 1,2$	p_{ij}	The initial (expected) processing time of job *j* on the machine *i*
j	Job index, $j = 1,2,\dots,n$	C_{ij}	The expected initial completion time of job *j* on the machine *i*
r	The expected value of repair times after each breakdown	C^r_{ij}	The expected real completion time of job *j* on the machine *i*

1. INTRODUCTION

The flow Shop Scheduling Problem (FSSP) covers many real case studies in practical problems [1]. Some papers described the applications of the two-machine flow shop scheduling problem (FSSP) [2–4]. Total completion time minimizes Work in Process (WIP) costs and the rapid turnaround of jobs. The two-machine FSSP with the sum of jobs' completion time as a primary objective, this paper's focus, even in deterministic scheduling environments, is strongly Np-hard [5]. Some effective heuristics proposed to cope with the problem's complexity do not seem superior over the other [6]. Besides, the job or machine-related uncertainties that lead to an interruption in the flow of jobs and result in unwanted delays are commonly occurring in the production environment, enhancing the problem's complexity. Arriving of an unanticipated new job [7], due date uncertainty [8], breakdown occurrence [9], uncertainty in job processing times [9, 10], etc are the

*Corresponding Author Institutional Email: *sahraeian@shahed.ac.ir* (R. Sahraeian)

likes of uncertainties and disruptions. In 70% of uncertainty oriented flow shop scheduling studies in past decades, the job processing time is uncertain, by 25%, the disruption is machine failure, and by 10%, both of these factors consider [11]. Machine failure and uncertain processing times discuss in this paper. Robust and stable scheduling is one of the policies in confronting uncertainty. The sensitivity of a schedule performance to its objective function is called robustness, but stability refers to the insensitivity of the start (or completion) time. Stability is a measure of changes in the sequence of jobs on a machine to the original. The concept of robustness is very close to flexibility: the ease of schedule reparability and the power of converting to new, high quality scheduling in the face of uncertainties. The expected realized total completion time has been implemented as a robustness measure by itself [8]. Here we take this definition as robustness. A function of the sum of deviation between each jobs' start/completion times in the initial and realized schedules are often the stability measure [8], and the same definition is accepted here. The value of the expected performance measure obtained by applying the righting shift policy on the initial schedule is a realized schedule. Additionally, a justified schedule with a small deviation from the initial one in the face of uncertainty and without significant degradation in the main objective is robust and stable. Simultaneous consideration of robustness and stability besides maintaining the schedule feasibility improves its flexibility against uncertainties.

Dealing with uncertainty-related deviations can be done with predictive, reactive, or predictive-reactive (hybrid) strategies [12]. In the predictive strategy, future uncertainties consider in the initial plan. Reactive or hybrid approaches are common strategies for dealing with machine breakdown. In almost all robustness-focused studies, dealing with machine failure disruption performs with reactive approaches or in the reactive phase of hybrid approaches [8]. In reactive strategies (e.g., rescheduling), especially in large-size problems, it takes a long time to deal with uncertainty. Predictive strategies can overcome this by actively preparing for any future uncertainties [12], so here we adopt a predictive approach to cope with machine breakdown and uncertainty of job processing times. The two-machine Flow Shop Scheduling Problem (*FSSP*) under uncertainty of processing time is referenced in many papers, commonly with makespan as a primary objective function [13, 14]. $C\max$, is also a primary objective in most *FSSP* under machine breakdown disruption studies [15, 16]. In addition, it has been a primary objective in the case of simultaneously considering the uncertainty of processing time and machine breakdown [9, 17]. Therefore, we define the robustness measure based on another performance measure; i.e., the total completion times. With a glimpse at the previous attempts in this realm of research, we can state contributions of this paper:

- Although various cases of the robust and stable flow shop-scheduling problem previously raised in studies, this article discusses a (particular) case of *F2* for the first time.
- Simultaneously considering robustness and stability to meet the requirements of producers and workers.
- Besides, predictively coping with the uncertainty of job processing times, dealing with machine breakdowns is also predictive.
- Proposing a novel robust and stable heuristics to cope with aforementioned-uncertainty conditions.
- The way of considering the uncertainty, the proposed solution method, and the primary objective function is different from the previous more related works.

The remainder of this paper is as follows; the related literature review is in section 2. In section 3, we define the problem and propose our solution method. In sections 4, 5, and 6, we presented computational results, managerial insight, and paper conclusions.

2. LITERATURE REVIEW

This paper presents robust and stable scheduling approaches for a permutation two-machine flow shop scheduling problem (PFSSP) under uncertainty with expected total completion times as a primary objective. According to the classification of Graham et al. [18], the problem denotes as $F2/prmu/\sum_{j=1}^{n} C_{2j}$, in the deterministic version, which is strongly Np-Hard [5]. The solution methods of $F2/prmu/\sum_{j=1}^{n} C_{2j}$ categorizing into exact and approximate methods. The high-performance problem-solving branch and bound algorithms and Lagrangian methods had proposed for $F2//\sum_{j=1}^{n} C_{2j}$ able to solve up to 50 jobs [19]. Due to the high complexity of this problem, heuristic methods have been ever the researchers' focus, for example, *MINITI* heuristic [6, 20]. Rossi et al. [21] proposed a simple high-efficiency FF-RN heuristic method by modifying the NEH heuristic and compared it with the best simple heuristics of PFSSP reviewed [21]. In the face of uncertainty, applying iterative simulation-based methods or producing robust (and stable) schedules are conventional approaches to encounter system disruptions [9]. Ghezail et al. [22] proposed a graphical robust, proactive approach to deal with uncertainty in the FSSP. Kasperski et al. [23] propose a predictive regret-based robust schedule with interval processing times. katrajeni et al. [24] propose a heuristic to minimize normalized makespan and instability in a dynamic flow shop under uncertainty of machine breakdown and job-ready time variability. Ying [25] applied Iterated Greedy (IG) and Simulated Annealing (SA) heuristics to produce a predictive regret-based robust schedule in a maximum completion time two-machine FSSP, whit interval processing times. Fazayeli et al. [26] applied the Genetic Algorithm (GA) and SA to produce a robust predictive schedule in an m-machine PFSP under uncertain repair time and machine breakdown. Rahmani [9] employed GA to propose a proactive-reactive robust, and stable schedule for a two-machine PFSP under uncertain job processing time and machine failure. Also, she applied scenarios to show the uncertainty of processing times, Cmax as an efficiency measure, maximum realized completion times of jobs as robustness, and the expected sum of square deviations between the completion time in actual and initial schedules stability measure. Cui et al. [16] used a simulation-based method to propose a robust predictive schedule for two-machine PFSP under machine breakdown with Cmax as an efficiency measure. Liao and Fu [7] exploit GA to propose a robust predictive schedule for an m-machine PFSP with interval processing times. Abtahi et al. [11] employed a robust optimization method to produce efficient, robust, and stable schedules in an m-machine FSSP under uncertainty. They applied scenarios to show the uncertainty of processing times, total completion times as an efficiency measure, total realized tardiness of jobs as robustness, and the expected sum of square deviations between the completion time in actual and initial schedules as a stability measure. Here we adopt total realize completion times as robustness and the expected sum of square deviations between the completion time in actual and initial schedules as a stability measure. We propose a modified Shifting bottleneck (*SB)* to produce robust partial solutions in a two-machine *FSSP* in the face of job processing times uncertainty and machine breakdown. Shifting bottleneck (*SB)* , a decomposition-based heuristic, performs well for job shops [27, 28]. Koulamas et al. [29] presented an efficient modified *SB* for two-machine PFSSP with total tardiness of jobs as a primary objective. Mukherjee et al. [30] showed that the modified *SB* is suitable in optimally solve a two-machine PFSSP with the makespan criterion. Elyasi and Salmasi [31] applied an adjusted *SB* in stochastic flow shop under due date uncertainty to minimize the number of tardy jobs. Allahverdi and Allahverdi [32] proposed a decomposition-based heuristics for a total completion time PFSSP with bounded processing time. As can be seen, few papers focused on

- Producing a robust and stable FSSP with total completion time as a primary objective.
- Predictively producing a robust and stable FSSP while considering the uncertainty of job processing time and machine breakdown simultaneously.

In this paper, we propose a heuristic method to produce a robust and stable schedule in a stochastic two-machine FSSP specified case with total completion time as a primary objective. Then we compare it to an exact solution method. The former (Our proposed heuristic) employs modified *SB* and a theorem of Abtahi et al. [33], and the latter uses a theorem of Pinedo [34] to hedge against job processing time uncertainty. We employ the Right-Shifting (RS) rescheduling method to obtain a realistic schedule after machine failures occurrence.

3. PROBLEM DEFINITION AND SOLUTION METHOD

In this paper, we considered a two-machine FSSP. The uncertain job processing time and random breakdowns of machines are the system disruptions. The processing time of job j on the first and second machines respectively follow the exponential distribution with rates λ_j and μ_j. The time between two consecutive failures follows an exponential distribution with the rate of θ and at most one failure is expected on a machine in each interval $(1/\theta)$. After each breakdown, minimal repairs perform to restore machines to the operating condition (which does not affect the machine age and breakdown parameter).

The following assumptions considered:

- All jobs are available at the beginning of the schedule,
- Machines have availability restriction; i.e., random machine break down may occur during the processing of job *j* on machine *i*,
- The time between two consecutive breakdowns follows an exponential distribution. Also, constant repair times allocate after each failure,
- The rest of the disrupted job will perform after machine repairing,
- Only non-delay schedules considered,
- The objective function is a minimization of schedules' robustness and stability simultaneously.

3. 1. Solution Method

According to a classification by Graham et al. [18], a problem of robust and stable two-machine FSSP under uncertainty of job processing time and machine breakdowns is represented as follows:

$$F2\left|\begin{matrix} \mathrm{p}_{1j} \sim \exp(\lambda_j), \mathrm{p}_{2j} \sim \exp(\mu_j); \\ brkdwn: U \sim \exp(\theta), D \sim G(t) \end{matrix}\right| \alpha.RM + (1-\alpha).SM$$

Pinedo [34] showed that sorting the jobs in descending order of $(\lambda_j - \mu_j)$ optimizes the expected total completion times (i.e., the intended robustness measure) in a two-machine FSSP particular case (when job processing times on the first (second) machine pursued the exponential distribution with the rate $\lambda_j(\mu_j)$). Here, we proposed two robust and stable methods, and for each one, two policies in the face of machine failure; reactive and predictive. To predictably deal with the machine failure, the buffer time insertion method, and to encounter with reactively, the right shift rescheduling (*RSH*) is implemented to the affected jobs [13] for details.

Two algorithms are proposed in this paper to handle such a problem:

- The optimal theorem based method (OBM).
- The decomposition-based method (DBM).

Based on a theorem, the OBM considering the uncertainty of processing times acquires the optimal robust solution (of Pinedo [34]) according to the decreasing order of $\lambda_j - \mu_j$.

The DBM method employs a modified shifting bottleneck heuristics and one-machine robustness and stability optimization theorem [33]; shifting bottleneck (SB) heuristics [5] decomposes a problem into sub-problems and solve each sub-problem optimally [25] using the shortest expected processing time (SEPT) first rule.

Theorem. SEPT rule solves

$$1\left|\begin{matrix} p_j \sim \exp(\lambda_j); \\ brkdwn: U \sim \exp(\theta), D \sim G_2(t) \end{matrix}\right| \alpha.RM + (1-\alpha).SM$$

optimally [33], where θ is the rate of machine breakdown and r is the expected repair time, $RM = E\sum_{j=1}^{n} C_j^{\ r}$ is a robustness measure, and $SM = E[\sum_{j=1}^{n}(C_j - C_j^{\ r})^2]$ is a stability measure. According to the above theorem, sorting the jobs in a no descending order of processing times over each machine seems an acceptable idea to give a robust and stable sequence for the intended uncertain two-machine FSSP. The steps of the DOM are as follows:

- Decompose the intended two-machine uncertain flow shop-scheduling problem into two one-machine sub-problems, with predefined conditions of uncertainty.
- Sequence the jobs according to the SEPT on the first machine.
- The first job on the first machine (M1) continued its process on the second machine regardless of the amount of its expected processing time (on M2).
- To determine the order of the remaining jobs on M2, do as follows.
- Whenever because of incomplete remaining (previous) jobs on M2, there is a queue with more than two jobs on M1, order the jobs queue according to the shortest expected value of their processing times on M2.
- Otherwise, the jobs keep their sequence on M1.

Figure 1 shows the flow chart of DBM. In the next section, we compared the proposed methods after the implementation of reactive as well as predictive policy. The job sequence on the two machines are the same in OBM; however, during the execution of DBM, the job sequence on *M1* may not keep.

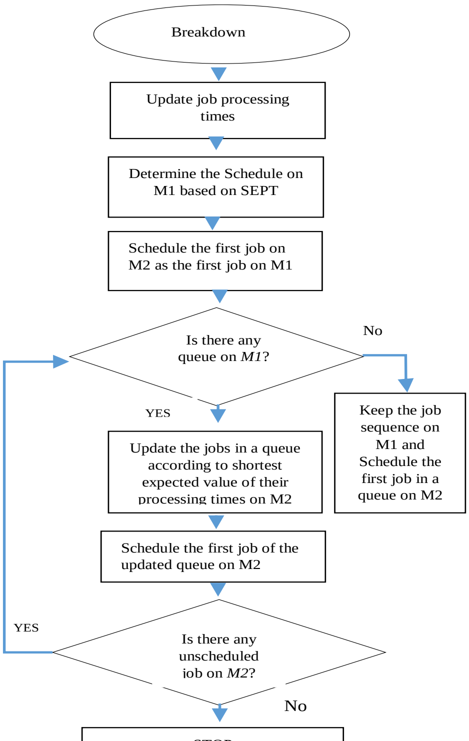


Figure 1. the flow chart of DBM

4. COMPUTATIONAL RESULTS

4. 1. Data Generation The job processing times on the first and second machines are uncertain and respectively follow the exponential distribution with the rate of λ_j and μ_j, where λ_j and μ_j are independently generated from $U[0.1, 1]$. We select the number of jobs from set $j = \{3, 5, 10, 30, 50, 60, 100\}$. Then 100 instances generate for each job number. Hence, we have 700 problems. For each test problem, we chose the rate of machine breakdown from a set $\theta = \{1/50, 1/60, 1/80\}$; a higher value of θ represents a higher probability of machine breakdown disruption.

Like Nouri et al. [35], the repair times duration follows an exponential distribution based on the meantime to repair value (*MTTR*) at two-level. The repair times' duration calculates via $r = \exp rnd(MTTR)$, and the *MTTR* based on the machine busy time (*MB*); for low level, $MTTR_l \in [0.01MB, 0.05MB]$, and for high level, $MTTR_h \in [0.05MB, 0.1MB]$. Ultimately we have combination of 4200 problems. The methods compare to reach a comprehensive conclusion as follows:

1. Without considering machine breakdown,
2. Applying the reactive policy after failure,
3. Dealing with breakdown disruption predictively.

4. 2. Two Comparative Methods without Considering Machine Breakdown Here, we examine the performance (objective function) of the two methods provide the managerial results for (the problem in question without assuming machine breakdown)

$F2/\mathrm{p}_{1j} \sim \exp(\lambda_j), \mathrm{p}_{2j} \sim \exp(\mu_j)/E\left(\sum_{j=1}^{n} C_{2j}^{r}\right)$. The problem coded in *MATLAB R2013b*, and the results have reported for different problems' sizes. Solving time is negligible and has not been brought. In Figures 2 to 4, *AEC*, *AEC0*, and *RD* respectively represent the Average Expected Completion time of the DBM to OBM and the relative deviation of the objective function of DBM to OBM without considering machine breakdowns. Figures 2 and 3 report the *AEC*, *AEC0*, and *RD* for the small size problems. It seems that DBM has a proper performance for small-size (3-30 jobs) problems (given that OBM offers the optimal solution for $F2/\mathrm{p}_{1j} \sim \exp(\lambda_j), \mathrm{p}_{2j} \sim \exp(\mu_j)/E\left(\sum_{j=1}^{n} C_{2j}^{r}\right)$). Although the performance of DBM as a heuristic method is still acceptable (see Figures 4 and 5), for the medium to large-size (50-100 jobs) problems, OBM thoroughly outperforms DBM.

4. 3. Two Comparative Methods Based on Applying Reactive Policy Here, we examine two proposed methods for (the problem in question)

$$F2\left|\begin{matrix}\mathrm{p}_{1j} \sim \exp(\lambda_j), \mathrm{p}_{2j} \sim \exp(\mu_j); \\ brkdwn: U \sim \exp(\theta), D \sim G(t)\end{matrix}\right|\alpha.RM + (1-\alpha).SM$$

applying the reactive policy after machine breakdown. The problem coded in *MATLAB R2013b*, and outputs has been reported different problems' sizes in Table 1. *RR_DBM*, *R_OBM*, *SR_DBM*, *SR_OBM*, *Z_RDBM*, *Z_ROBM*, *n*, and *TIME*, respectively represent the robustness of *DBM* and *OBM*, the stability of *DBM* and *OBM*, the objective function value of *DBM* and *OBM*, the number of jobs, and the problem-solving time by applying reactive policy after machine breakdown. Figures 6 to 10 illustrate the contents of Table 1. According to Figures 6 to 10, in all cases of the small-size problem (n<=10), regardless of the values of *TETA* and *MTTR*; there is no significant difference between the two proposed methods' performance. Nevertheless, in medium to large-size problems (n>=30), *OBM* outperforms *DBM* by applying reactive policy after machine breakdown. These results had obtained by considering the same coefficients for robustness and stability ($\alpha = (1-\alpha) = 0.5$).

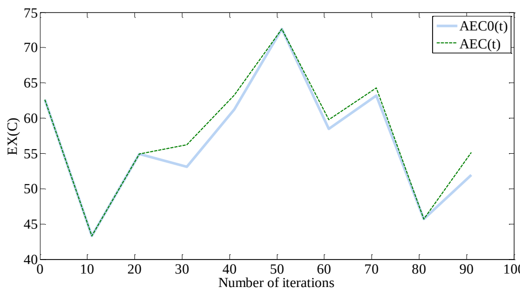


Figure 2. The comparison the expected completion time for two methods without considering machine breakdown (small-size problem)

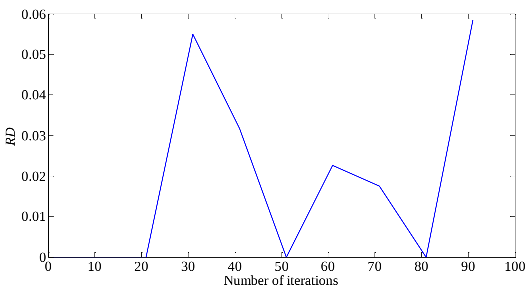


Figure 3. The related deviation between two methods without considering machine breakdown (small-size problem)

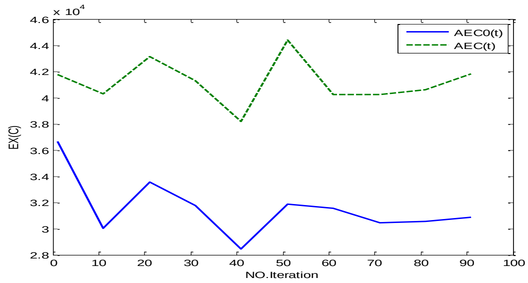


Figure 4. Comparing the two methods' expected completion time without considering machine breakdown (medium to large-size problems)

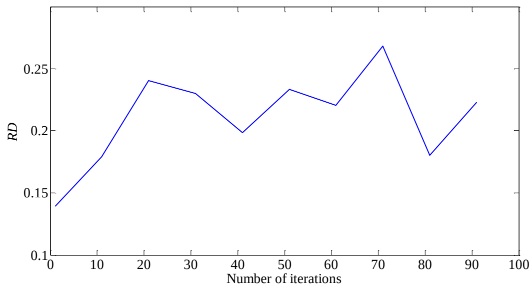


Figure 5. The related deviation between two methods without considering machine breakdown (medium to large-size problems)

4. 4. Two Comparative Methods' Based on Applying Predictive Policy Here, we examine the two proposed methods for solving (the problem in question)

$$F2\left|\begin{matrix}\mathrm{p}_{1j} \sim \exp(\lambda_j), \mathrm{p}_{2j} \sim \exp(\mu_j); \\ brkdwn: U \sim \exp(\theta), D \sim G(t)\end{matrix}\right|\alpha.RM + (1-\alpha).SM$$

by applying the predictive policy to encounter the machine breakdown. The problem coded in *MATLAB R2013b*, and the results were reported for different problems' sizes in Table 2 and Figure 11. *RP_DBM*, *RP_OBM*, *SP_DBM*, *SP_OBM*, *Z_PDBM*, *ZP_OBM*, *n*, and *TIME*, respectively represent the robustness of *DBM*

TABLE 1. Robustness, stability, and the objective function of the two methods by applying the reactive policy for different problem-parameters

No	1/TETA	MTTR	n	RR_DBM	RR_OBM	SR_DBM	SR_OBM	Z_RDBM	Z_ROBM	T
1	80	low	3	0.01	0.05	0.05	0.02	0.03	0.036	1.58
2	80	low	5	0.3	0.23	0.26	0.1	0.28	0.16	1.8
3	80	low	10	0.27	0.52	0.47	0.49	0.36	0.5	6.48
4	80	low	30	21.7	1.8	6.7	5.6	14.2	3.7	99
5	80	low	50	237	3.5	43	23.8	136	7.52	308
6	80	low	60	335	5	63	77	189	14.5	648
7	80	low	100	5858.3	15.3	226.7	79	3042.5	47	2798
8	60	low	3	0.018	0.001	0.02	0.001	0.01	0.009	27
9	60	low	5	0.05	0.33	0.18	0.12	0.11	0.22	11
10	60	low	10	0.82	0.32	0.67	0.19	0.75	0.26	11
11	60	low	30	23	4	7.7	8.3	15.5	6	202
12	60	low	50	574	4	44	15	310	10	608
13	60	low	60	1733	4	73	12	903	9	380
14	60	low	100	4279	15	223	119	1251	68	3187
15	50	low	3	0.02	0.1	0.08	0.04	0.05	0.08	1.38
16	50	low	5	0.49	0.18	0.23	0.07	0.36	0.12	1.94
17	50	low	10	0.62	0.54	0.72	0.46	0.67	0.5	5.43
18	50	low	30	18	2	7.5	3.4	12.8	2.8	51.3
19	50	low	50	475	4	45	13	260	9	213.2
20	50	low	60	1581	7.5	81	50	831	29	363
21	50	low	100	4013	13	245	93	2128	54	1567
22	80	high	3	0.07	0.17	0.11	0.17	0.09	0.17	1.4
23	80	high	5	0.43	0.4	0.4	0.41	0.42	0.41	1.8
24	80	high	10	3	1.08	1.56	1.48	2.28	1.28	6.3
25	80	high	30	147	5	23	17	85.3	11.16	72.3
26	80	high	50	42.5	46	329	8	186	27	294
27	80	high	60	2663	27	125	343	1394	185	455
28	80	high	100	12388	94.5	418	2055	6403	1075	2049
29	60	high	3	0.21	0.21	0.17	0.24	0.19	0.23	1.41
30	60	high	5	1.03	0.42	0.46	0.50	0.75	0.46	1.62
31	60	high	10	1.18	1.9	1.1	2.70	1.5	1.9	4.6
32	60	high	30	166.9	6.6	25.7	42.8	96.3	6.6	70.4
33	60	high	50	378.7	7.8	50.7	58.5	219.2	33.2	313.7
34	60	high	60	1863.8	16.2	212.5	113.4	988.6	114.3	511.6
35	60	high	100	11129.2	40.65	402	543	5765.7	291.8	2947
36	50	high	3	0.09	0.16	0.13	0.08	0.11	0.12	1.5
37	50	high	5	0.38	0.39	0.35	0.35	0.367	0.37	2.11
38	50	high	10	5.46	1.09	1.9	1.5	3.7	1.3	5.27
39	50	high	30	103.6	12.27	17.1	12.27	60.36	8.2	53.98
40	50	high	50	1384.3	19.8	97.6	170.3	740.99	95	250
41	50	high	60	2708.8	14.9	146.65	151.99	1427.7	83.47	444.9
42	50	high	100	29.6	6328	441.2	316.9	3384.7	173.2	3017

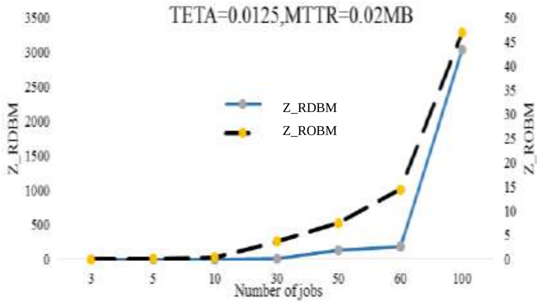


Figure 6. Comparing the objective function of the two methods by applying the reactive policy, Low expected failure, and short repair time

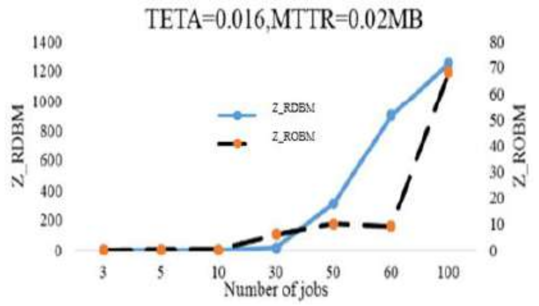


Figure 7. Comparing the objective function of the two methods by applying the reactive policy, medium expected failure, and short repair time

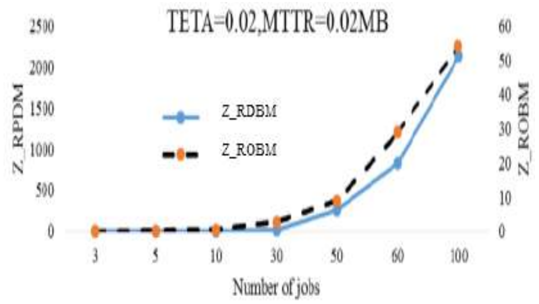


Figure 8. Comparing the objective function of the two methods by applying the reactive policy, low expected failure, and short repair time

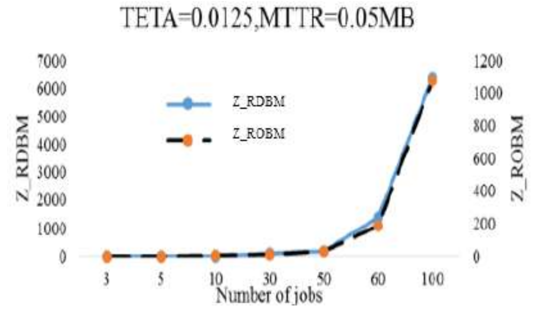


Figure 9. Comparing the objective function of the two methods by applying the reactive policy, low expected failure, and high repair time

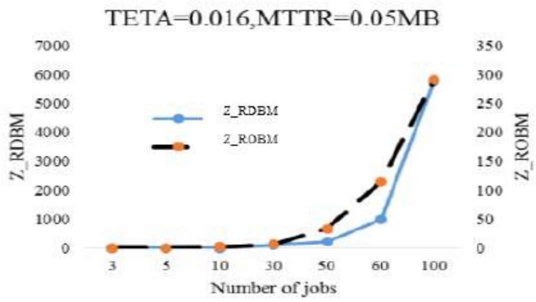


Figure 10. Comparing the objective function of the two methods by applying the reactive policy, medium expected failure, and high repair time

and *OBM*. The stability of *DBM* and *OBM*, the objective function value of the *DBM* and *OBM*, the number of jobs and the problem-solving time incorporated by applying predictive policy to encounter with machine breakdown. Figure 11 illustrates the contents of Table 2. According to Figure 11 and Table 2, DBM is preferred to OBM in all cases regardless of the values of TETA, and MTTR, especially when the number of jobs increases. These results were obtained by considering the same coefficients for robustness and stability ($\alpha = (1-\alpha) = 0.5$). In the sensitivity analysis section, we will analyze the effect of different values of α (robustness coefficient) on the performance measure of the two methods.

4. 4. Sensitivity Analysis This section provides additional tests on the methods' parameters to gauge their effects on the objective functions' values.

4. 4. 1. Testing on the Rate of Machine Breakdown and Mean Time to Repair According to Tables 3 to 8 and Figures 12 to 14, regardless of the number of jobs, the rate of a machine breakdown and the meantime to repair, DBM outperforms OBM. Also, as expected, the higher the failure rate (TETA) and the meantime to repair (MTTR), the worse the value of the robustness, stability, and two methods' objective functions.

4. 4. 2. Testing on the Stability and Robustness Coefficients In this section, different values of the robustness coefficients (α) had applied to achieve both methods' objective values. The results depicted for the low level of *MTTR* and $\theta = 0.0125$ in Figures 15 to 18. The effects of the varying coefficients of the robustness on the two methods' objective functions by applying the predictive policy showed in Table 9. According to Table 9 and Figure 17, OBM outperforms DBM when $\alpha \geq 0.7$, especially for many jobs. For values less than 0.7, DBM is superior to the OBM. In comparing two methods by applying the reactive policy, the effect of the robustness coefficient ignores. In this case, OBM is always almost outperformed DBM.

TABLE 2. Robustness, stability, and two methods' objective function by applying the predictive policy for different problem parameters

No	1/TETA	MTTR	n	RP_DBM	RP_OBM	SP_DBM	SP_OBM	ZP_DBM	ZP_OBM	T
1	80	low	3	0.01	0.069	0.067	0.015	0.039	0.042	1.58
2	80	low	5	0.27	0.2	0.27	0.07	0.27	0.13	1.8
3	80	low	10	0.26	0.95	0.73	0.32	0.5	0.48	6.48
4	80	low	30	16.6	16.6	9.5	23.7	13	20	99
5	80	low	50	221.25	85.23	44	291.7	133	188.5	308
6	80	low	60	524	164	73	1313	299	739	648
7	80	low	100	6637	667	366	9029	3501	4848	2798
8	60	low	3	0.002	0.05	0.05	0.002	0.25	0.25	27
9	60	low	5	0.04	0.26	0.25	0.08	0.17	0.15	11
10	60	low	10	0.54	0.36	0.55	0.13	0.54	0.25	11
11	60	low	30	19.5	28	13	64	16	46	202
12	60	low	50	552	134	68	936	309	535	608
13	60	low	60	1514	213	110	2086	812	1149	380
14	60	low	100	14350	1018	529	22181	7640	11600	3187
15	50	low	3	0.02	0.1	0.1	0.03	0.06	0.07	1.38
16	50	low	5	0.5	0.2	0.3	0.04	0.4	0.1	1.94
17	50	low	10	0.4	1.1	0.9	0.4	0.65	0.77	5.43
18	50	low	30	21	26	12.5	52	17	39	51.3
19	50	low	50	588	146	73.5	972	331	559	213.2
20	50	low	60	2280	265	149	1361	1214	1713	363
21	50	low	100	1241	20487	608	41905	10548	21573	1567
22	80	high	3	0.11	0.2	0.19	0.14	0.15	0.17	1.4
23	80	high	5	0.35	0.36	0.41	0.29	0.38	0.33	1.8
24	80	high	10	2.8	1.28	1.8	1.01	2.3	1.1	6.3
25	80	high	30	74.8	29	19.9	96.5	47.4	49.25	72.3
26	80	high	50	490	167	76	1399.8	283	783	294
27	80	high	60	4458	384	208.8	5710	2333.5	3047	455
28	80	high	100	30377.5	1569	705	61743	15541	31656	2049
29	60	high	3	0.16	0.21	0.2	0.18	0.18	0.2	1.4
30	60	high	5	0.91	0.38	0.51	0.29	0.71	0.34	1.62
31	60	high	10	2	1.9	1.8	1.9	1.9	1.9	4.6
32	60	high	30	159.6	54.2	28.1	244.5	93.8	149.4	70.4
33	60	high	50	973.7	202.1	96.1	2032.4	535	1117.3	313.7
34	60	high	60	2248.8	416	167.2	7344.1	1208	3880	511.6
35	60	high	100	39551	1658	815	76047	20183	38852	2948
36	50	high	3	0.06	0.15	0.14	0.05	0.1	0.1	1.5
37	50	high	5	0.26	0.38	0.37	0.22	0.32	0.30	2.11
38	50	high	10	8.6	2.16	2.46	2	5.5	2.08	5.27
39	50	high	30	201	58.87	33.97	312.85	117.5	185.86	53.98
40	50	high	50	4492.4	176.9	5718.1	95.01	2334.68	3029.5	250
41	50	high	60	6660	523.8	264.3	12592	3462.3	6558	444.9
42	50	high	100	61755.5	2149.5	1119.6	110545.6	31437.6	56347.5	3017

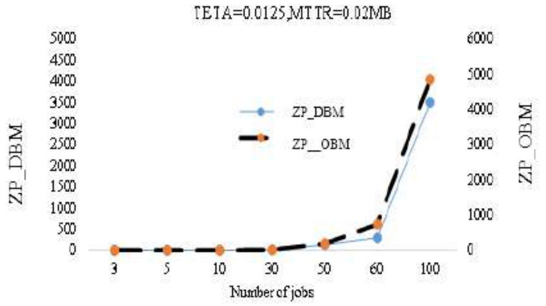


Figure 11. Comparison of two methods' objective function by applying the predictive policy, low expected failure, and short repair time

TABLE 3. Robustness, stability, and the objective function of two methods by applying the predictive policy for n=100, short MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
100	0.0125	4443	2374	8202	401	684	4293
MTTR	0.016	11600	7640	22181	529	1018	14350
0.02MB	0.02	21573	10548	41905	608	20487	1241

TABLE 4. Robustness, stability, and the objective function of two methods by applying the predictive policy for n=100, high MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
100	0.0125	31656	15541	61743	705	1569	30373
MTTR	0.016	38852	20183	76047	815	1658	39551
0.05MB	0.02	56347	31437	110545	1119	2149	61755

TABLE 5. Robustness, stability, and the objective function for the two methods by applying the predictive policy for n=50, low MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
50	0.0125	359	191	300	82	57	662
MTTR	0.016	535	309	936	68	134	552
0.02MB	0.02	559	331	972	73.5	146	588

TABLE 6. Robustness, stability, and the objective function of the two methods by applying the predictive policy for n=50, high MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
50	0.0125	783	283	1399	76	167	490
MTTR	0.016	1117	535	2032	96	202	973
0.05MB	0.02	3029	2334	95	571	177	4492

TABLE 7. Robustness, stability, and the objective function of the two methods applying the predictive policy for n=30, low MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
100	0.0125	30	21	31	12.7	21.7	29
MTTR	0.016	46	16	64	13	28	19.5
0.02MB	0.02	39	17	52	12.5	26	21

TABLE 8. Robustness, stability, and the objective function of the two methods by applying the predictive policy for n=30, high MTTR, and different values of TETA

n	TETA	ZP_ OBM	ZP_ DBM	SP_ OBM	SP_ DBM	RP_ OBM	RP_ DBM
100	0.0125	49.25	47.4	96.5	19.9	29	74.8
MTTR	0.016	149.4	93.8	244.5	28.1	54.2	159.6
0.05MB	0.02	185.9	117.5	312.8	33.97	58.87	201

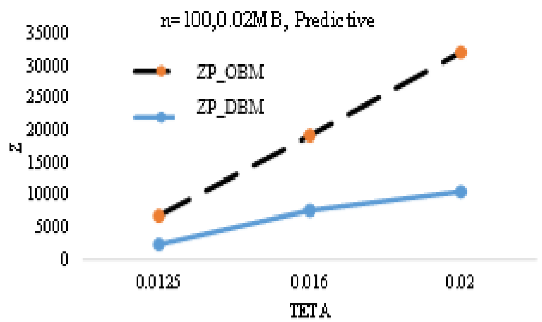


Figure 12. Comparison of two methods' objective function by applying the predictive policy for n=100, high MTTR, and different values of TETA

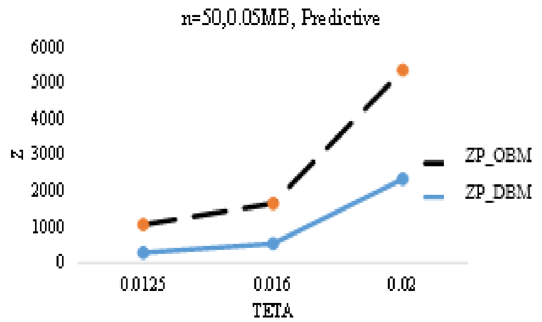


Figure 13. Comparison of two methods' objective function applying the predictive policy for n=50, high MTTR, and different values of TETA

5. MANAGERIAL INSIGHT

By increasing the importance of producer satisfaction level (robustness) to the satisfaction level of the production environment (stability), i.e., $\alpha \geq 0.7$, OBM outperforms DBM. In other words, if the production system's interior completely aligns with the producer's goals and the producer is not worried about the reaction of the production staff, select OBM, and otherwise DBM.

If $\theta < 0.0125$, OBM outperforms DBM, i.e., if the wear of the machines is negligible, and the predictive method has a higher cost than the reactive, select OBM, and otherwise DBM.

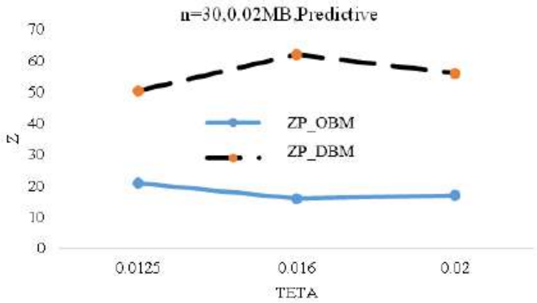


Figure 14. Comparison of two methods' objective function by applying the predictive policy for n=30, low MTTR, and different values of TETA

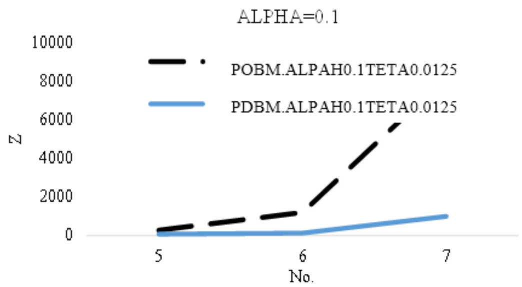


Figure 15. Comparison of two methods' objective function by applying the predictive policy for ALPHA=0.1, low level of MTTR, TETA

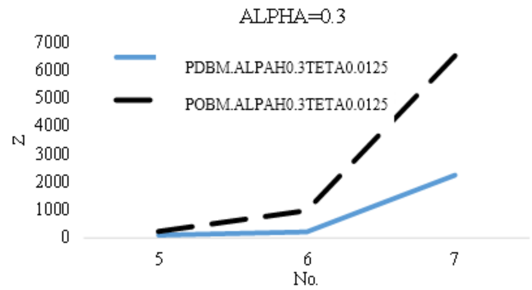


Figure 16. Comparison of two methods' objective function by applying the predictive policy, ALPHA=0.3, low level of MTTR, TETA

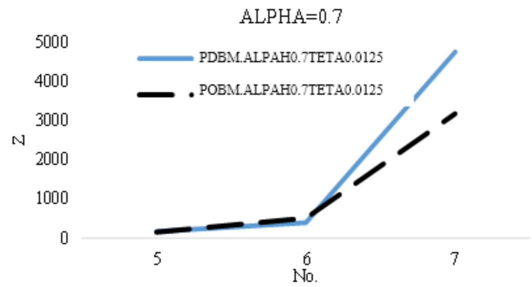


Figure 17. Comparison of two methods' objective function by applying the predictive policy, ALPHA=0.7, low level of MTTR, TETA

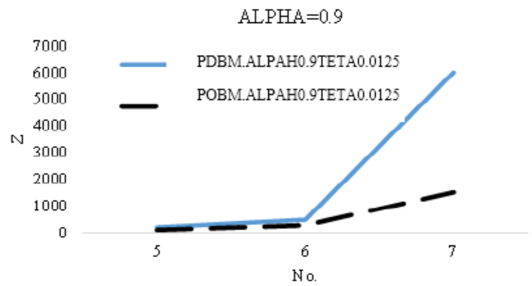


Figure 18. Comparison of two methods' objective function by applying the predictive policy, ALPHA=0.9, low level of MTTR, TETA

TABLE 9. Comparison of two methods' objective function applying the predictive policy for different values of the robustness coefficient

α	0.1	0.1	0.3	0.3	0.7	0.7	0.9	0.9
No	**Z_DBM**	**ZP_OBM**	**Z_DBM**	**ZP_OBM**	**Z_DBM**	**ZP_OBM**	**Z_DBM**	**ZP_OBM**
1	0.0613	0.0204	0.0499	0.0312	0.0271	0.0528	0.0157	0.064
2	0.27	0.083	0.27	0.109	0.27	0.161	0.27	0.187
3	0.683	0.383	0.589	0.509	0.401	0.761	0.307	0.887
4	10.21	22.99	11.63	21.57	14.47	18.73	15.89	17.31
5	61.725	271.05	97.175	229.75	168.07	147.17	203.5	105.9
6	118.1	1198.1	208.3	968.3	388.7	508.7	478.9	278.9
7	993.1	8192.8	2247.3	6520.4	4755.7	3175.6	6009.9	1503
8	0.0452	0.0068	0.0356	0.0164	0.0164	0.0356	0.0068	0.045
9	0.229	0.098	0.187	0.134	0.103	0.206	0.061	0.242

10	0.549	0.153	0.547	0.199	0.543	0.291	0.541	0.337
11	13.65	60.4	14.95	53.2	17.55	38.8	18.85	31.6
12	116.4	855.8	213.2	695.4	406.8	374.6	503.6	214
13	250.4	1898.7	531.2	1524.1	1092.8	774.9	1373.6	400
14	1911.1	20064	4675.3	15832	10204	7366.9	12968	3134
15	0.092	0.037	0.076	0.051	0.044	0.079	0.028	0.093
16	0.32	0.056	0.36	0.088	0.44	0.152	0.48	0.184
17	0.85	0.47	0.75	0.61	0.55	0.89	0.45	1.03
18	13.35	49.4	15.05	44.2	18.45	33.8	20.15	28.6
19	124.95	889.4	227.85	724.2	433.65	393.8	536.55	228.6
20	362.1	1251.4	788.3	1032.2	1640.7	593.8	2066.9	374.6
21	671.3	39763	797.9	35480	1051.1	26912	1177.7	22629
22	0.182	0.146	0.166	0.158	0.134	0.182	0.118	0.194
23	0.404	0.297	0.392	0.311	0.368	0.339	0.356	0.353
24	1.9	1.037	2.1	1.091	2.5	1.199	2.7	1.253
25	25.39	89.75	36.37	76.25	58.33	49.25	69.31	35.75
26	117.4	1276.5	200.2	1029.9	365.8	536.84	448.6	290.3
27	633.72	5177.4	1483.5	4112.2	3183.2	1981.8	4033	916.6
28	3672.2	55726	9606.7	43691	21476	19621	2741	7586
29	0.196	0.183	0.188	0.189	0.172	0.201	0.164	0.207
30	0.55	0.299	0.63	0.317	0.79	0.353	0.87	0.371
31	1.82	1.9	1.86	1.9	1.94	1.9	1.98	1.9
32	41.25	225.47	67.55	187.41	120.15	111.29	146.45	73.23
33	183.86	1849.3	359.38	1483.3	710.42	751.19	885.94	385.13
34	375.36	6651.2	791.68	5265.6	1624.3	2494.4	2040.	1108.
35	4688.6	68608	12436	53730.	27930.	23974.	35677	9097
36	0.132	0.06	0.116	0.08	0.084	0.12	0.068	0.14
37	0.359	0.236	0.337	0.268	0.293	0.332	0.271	0.364
38	3.074	2.016	4.302	2.048	6.758	2.112	7.986	2.144
39	50.673	287.45	84.079	236.65	150.8	135.06	184.3	84.3
40	5595.5	103.2	5350.4	119.6	4860.1	152.33	4615	168.7
41	903.87	11385.	2183.0	8971.5	4741.3	4144.3	6020.4	1730.
42	7183.1	99706	19310	78027	43565	34668	55692	12989

6. CONCLUSION

In this study, we simultaneously considered the uncertainty of processing time and machine breakdowns in a two-machine flow shop scheduling problem. Two methods were proposed and compared in three situations to deal with this problem; without considering machine failure disruption and considering machine breakdown applying with the reactive and predictive policy. In the first situation, decomposition-based methods have an acceptable performance compared with the optimal base one. In the second status, OBM had a higher performance than DBM except in small-size problems. In applying the predictive policy, DBM had a higher performance than OBM, except in cases where producer satisfaction is more important than stability in the production environment. In all considering situations, the problem-solving time was acceptable and almost close to each other. Finally, OBM applying with the reactive policy, due to its lower objective function and its lower cost to DBM, seems more appropriate to solve the problem.

In this paper, a general approach proposed that can be used for robustness and stability optimization in an m-machine flow shop or job shop scheduling problem, with other measures of robustness and stability, or in the construction of predictive-reactive methods.

7. REFERENCES

1. Fuchigami, H. Y., and Rangel, S., "A survey of case studies in production scheduling: Analysis and perspectives", ***Journal of Computational Science***, Vol. 25, (2018), 425–436. doi:10.1016/j.jocs.2017.06.004
2. Mokhtari, H., Molla-Alizadeh, S., and Noroozi, A., "Modelling and Optimization of Reliability of a Flowshop Production System via Preventive Maintenance Policy", ***International Journal of Engineering, Transaction C: Aspects***, Vol. 28, No. 12, (2015), 1774–1781. doi:10.5829/idosi.ije.2015.28.12c.10
3. Yazdi, M., Zandieh, M., and Haleh, H., "A Mathematical Model for Scheduling Elective Surgeries for Minimizing the Waiting Times in Emergency Surgeries", ***International Journal of Engineering, Transaction C: Aspects***, Vol. 33, No. 3, (2020), 448–458. doi:10.5829/ije.2020.33.03c.09
4. Abtahi, Z., Sahraeian, R., and Rahmani, D., "A Stochastic Model for Prioritized Outpatient Scheduling in a Radiology Center", ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 4, (2020), 598–606. doi:10.5829/ije.2020.33.04a.11
5. Pinedo, M. L., Scheduling, (2012), Boston, MA, New York: Springer. doi:10.1007/978-1-4614-2361-4
6. Gupta, J. N. D., and Stafford, E. F., "Flowshop scheduling research after five decades", ***European Journal of Operational Research***, Vol. 169, No. 3, (2006), 699–711. doi:10.1016/j.ejor.2005.02.001
7. Liao, W., and Fu, Y., "Min–max regret criterion-based robust model for the permutation flow-shop scheduling problem", ***Engineering Optimization***, Vol. 52, No. 4, (2020), 687–700. doi:10.1080/0305215X.2019.1607848
8. Drwal, M., and Józefczyk, J., "Robust min–max regret scheduling to minimize the weighted number of late jobs with interval processing times", ***Annals of Operations Research***, Vol. 284, No. 1, (2020), 263–282. doi:10.1007/s10479-019-03263-6
9. Rahmani, D., "A new proactive-reactive approach to hedge against uncertain processing times and unexpected machine failures in the two-machine flow shop scheduling problems", ***Scientia Iranica***, Vol. 24, No. 3, (2017), 1571–1584. doi:10.24200/sci.2017.4136
10. Framinan, J. M., Fernandez-Viagas, V., and Perez-Gonzalez, P., "Using real-time information to reschedule jobs in a flowshop with variable processing times", ***Computers & Industrial Engineering***, Vol. 129, (2019), 113–125. doi:10.1016/j.cie.2019.01.036
11. Abtahi, Z., Sahraeian, R., and Rahmani, D., "A New Approach Generating Robust and Stable Schedules in m-Machine Flow Shop Scheduling Problems: A Case Study", ***International Journal of Engineering, Transaction B: Applications***, Vol. 33, No. 2, (2020), 293–303. doi:10.5829/ije.2020.33.02b.14
12. Mehta, S. V., and Uzsoy, R. M., "Predictable scheduling of a job shop subject to breakdowns", ***IEEE Transactions on Robotics and Automation***, Vol. 14, No. 3, (1998), 365–378. doi:10.1109/70.678447
13. Lee, J.-Y., and Kim, Y.-D., "Minimizing total tardiness in a two-machine flowshop scheduling problem with availability constraint on the first machine", ***Computers & Industrial Engineering***, Vol. 114, (2017), 22–30. doi:10.1016/j.cie.2017.10.004
14. Ma, S., Wang, Y., and Li, M., "A Novel Artificial Bee Colony Algorithm for Robust Permutation Flowshop Scheduling", In: Natural Computing for Unsupervised Learning, (2019), 163–182, Springer, Cham. doi:10.1007/978-3-319-98566-4_8
15. Liu, F., Wang, S., Hong, Y., and Yue, X., "On the Robust and Stable Flowshop Scheduling Under Stochastic and Dynamic Disruptions", ***IEEE Transactions on Engineering Management***, Vol. 64, No. 4, (2017), 539–553. doi:10.1109/TEM.2017.2712611
16. Cui, W., Lu, Z., Li, C., and Han, X., "A proactive approach to solve integrated production scheduling and maintenance planning problem in flow shops", ***Computers & Industrial Engineering***, Vol. 115, (2018), 342–353. doi:10.1016/j.cie.2017.11.020
17. Shen, J., and Zhu, Y., "Uncertain flexible flow shop scheduling problem subject to breakdowns", ***Journal of Intelligent & Fuzzy Systems***, Vol. 32, No. 1, (2017), 207–214. doi:10.3233/JIFS-151400
18. Graham, R. L., Lawler, E. L., Lenstra, J. K., and Kan, A. H. G. R., "Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey", In: Annals of Discrete Mathematics (Vol. 5), (1979), 287–326, Elsevier. doi:10.1016/S0167-5060(08)70356-X
19. Abedinnia, H., Glock, C. H., and Brill, A., "New simple constructive heuristic algorithms for minimizing total flow-time in the permutation flowshop scheduling problem", ***Computers & Operations Research***, Vol. 74, (2016), 165–174. doi:10.1016/j.cor.2016.04.007
20. Rakrouki, M. A., Kooli, A., Chalghoumi, S., and Ladhari, T., "A branch-and-bound algorithm for the two-machine total completion time flowshop problem subject to release dates", ***Operational Research***, Vol. 20, No. 1, (2020), 21–35. doi:10.1007/s12351-017-0308-7
21. Rossi, F. L., Nagano, M. S., and Sagawa, J. K., "An effective constructive heuristic for permutation flow shop scheduling problem with total flow time criterion", ***The International Journal of Advanced Manufacturing Technology***, Vol. 90, Nos. 1–4, (2017), 93–107. doi:10.1007/s00170-016-9347-0
22. Ghezail, F., Pierreval, H., and Hajri-Gabouj, S., "Analysis of robustness in proactive scheduling: A graphical approach", ***Computers & Industrial Engineering***, Vol. 58, No. 2, (2010), 193–198. doi:10.1016/j.cie.2009.03.004
23. Kasperski, A., Kurpisz, A., and Zieliński, P., "Approximating a two-machine flow shop scheduling under discrete scenario uncertainty", ***European Journal of Operational Research***, Vol. 217, No. 1, (2012), 36–43. doi:10.1016/j.ejor.2011.08.029
24. Katragjini, K., Vallada, E., and Ruiz, R., "Flow shop rescheduling under different types of disruption", ***International Journal of Production Research***, Vol. 51, No. 3, (2013), 780–797. doi:10.1080/00207543.2012.666856
25. Ying, K.-C., "Scheduling the two-machine flowshop to hedge against processing time uncertainty", ***Journal of the Operational Research Society***, Vol. 66, No. 9, (2015), 1413–1425. doi:10.1057/jors.2014.100
26. Fazayeli, M., Aleagha, M.-R., Bashirzadeh, R., and Shafaei, R., "A hybrid meta-heuristic algorithm for flowshop robust scheduling under machine breakdown uncertainty", ***International Journal of Computer Integrated Manufacturing***, Vol. 29, No. 7, (2016), 709–719. doi:10.1080/0951192X.2015.1067907
27. Pinedo, M., and Singer, M., "A shifting bottleneck heuristic for minimizing the total weighted tardiness in a job shop", ***Naval Research Logistics (NRL)***, Vol. 46, No. 1, (1999), 1–17. doi: https://doi.org/10.1002/(SICI)1520-6750(199902)46:1<1::AID-NAV1>3.0.CO;2-%23
28. Shi, F., Zhao, S., and Meng, Y., "Hybrid algorithm based on improved extended shifting bottleneck procedure and GA for assembly job shop scheduling problem", ***International Journal of Production Research***, Vol. 58, No. 9, (2020), 2604–2625. doi:10.1080/00207543.2019.1622052
29. Koulamas, C., "A guaranteed accuracy shifting bottleneck algorithm for the two-machine flowshop total tardiness problem", ***Computers & Operations Research***, Vol. 25, No. 2, (1998), 83–89. doi:10.1016/S0305-0548(97)00028-2
30. Mukherjee, S., and Chatterjee, A. K., "Applying machine based

decomposition in 2-machine flow shops", ***European Journal of Operational Research***, Vol. 169, No. 3, (2006), 723–741. doi:10.1016/j.ejor.2004.10.028

31. Elyasi, A., and Salmasi, N., "Stochastic flow-shop scheduling with minimizing the expected number of tardy jobs", ***The International Journal of Advanced Manufacturing Technology***, Vol. 66, Nos. 1–4, (2013), 337–346. doi:10.1007/s00170-012-4328-4
32. Allahverdi, M., and Allahverdi, A., "Minimizing total completion time in a two-machine no-wait flowshop with uncertain and bounded setup times", ***Journal of Industrial & Management Optimization***, Vol. 16, No. 5, (2020), 2439–2457. doi:10.3934/jimo.2019062
33. Abtahi, Z., Sahraeian, R., and Rahmani, D., "Predictive heuristics to generate robust and stable schedules in single machine systems under disruptions", ***Scientia Iranica***, Vol. 16, No. 5, (2019), 0–0. doi:10.24200/sci.2019.50809.1873
34. Pinedo, M., "Optimal policies in stochastic shop scheduling", ***Annals of Operations Research***, Vol. 1, No. 3, (1984), 305–329. doi:10.1007/BF01874395
35. Nouiri, M., Bekrar, A., Jemai, A., Trentesaux, D., Ammari, A. C., and Niar, S., "Two stage particle swarm optimization to solve the flexible job shop predictive scheduling problem considering possible machine breakdowns", ***Computers & Industrial Engineering***, Vol. 112, (2017), 595–606. doi:10.1016/j.cie.2017.03.006

Persian Abstract

چکیده

در این مقاله یک رویکرد پیش‌بینانه‌ی مقاوم و پایدار برای مسأله‌ی زمان‌بندی جریان کارگاهی دو ماشینی با فرض زمان فرآیند احتمالی کارها و اختلال خرابی ماشین ارائه شده است. در واقع، یک روش کلی ارائه می‌شود که می‌توان از آن برای بهینه‌سازی مقاوم و پایدار در مسأله زمان‌بندی محیط کارگاهی و تولید محصول استفاده کرد. مقیاس مقاومت، مقدار مورد انتظار مجموع زمان‌های اتمام کارها در زمان‌بندی واقعی است. مقیاس پایداری، مقدار مورد انتظار مربع مجموع انحرافات زمان اتمام کارها در برنامه‌ریزی اولیه و واقعی است. به منظور حل این مسأله، دو روش پیشنهاد شده و مورد مقایسه قرار گرفته است. یک روش مبتنی بر تجزیه مسأله مورد نظر به دو زیر مسأله و حل هر زیرمسأله به صورت بهینه و روش دیگر، بر پایه‌ی بر یک قضیه‌ی ریاضی بر اساس نتایج محاسباتی است. روش دوم از نظر مقاومت و پایداری، به ویژه در مورد مسائل با اندازه‌ی بزرگ، عملکرد بهتری دارد. به عبارت دیگر، روش دوم به دلیل بهبود قدرت پاسخگویی تولید کننده به مشتری و افزایش رضایت کارکنان خط تولید، برای حل این مسأله پیشنهاد می‌شود.

IJE TRANSACTIONS A: Basics Vol. 34, No. 4, (April 2021) 948-955

International Journal of Engineering

Journal Homepage: www.ije.ir

Utilization of Gene Expression Programming for Modeling of Mechanical Performance of Titanium/Carbonated Hydroxyapatite Nanobiocomposites: The Combination of Artificial Intelligence and Material Science

M. R. Shojaei[a], G. R. Khayati*[b], A. Hasani[a]

[a] *Department of Materials Science and Engineering, Sharif University of Technology, Tehran, Iran*
[b] *Department of materials science and Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*

PAPER INFO

Paper history:
Received 23 November 2020
Received in revised form 22 December 2020
Accepted 16 January 2021

Keywords:
Ti/CHA Nanocomposite
Mechanical Alloying
Powder Metallurgy
Biomaterials
GEP

ABSTRACT

Titanium carbonated hydroxyapatite (Ti/CHA) nanobiocomposites have extensive biological applications due to the excellent biocompatibility and similar characteristics to the human bone. Ti/CHA nanobiocomposite has good biological properties but it suffer from diverse characteristics especially in hardness, Young's modulus, apparent porosity and relative density. This investigation is an attempt to propose the predictive models using gene expression programming (GEP) to estimate these characteristics. In this regards, GEP is used to model and compare the effect of practical variables including pressure, Ti/CHA contents and sintering temperature on their monitored properties. To achieve this goal, 90 different experiments were considered to create the GEP models. Selected data set were divided randomly into 63 training sets and 27 testing sets. Finally, five of the best models were reported for each different output. Sensitivity analyses were done to determine and rank the practical parameters on each of the investigated properties and revealed that wt.% Ti, wt.% CHA, compaction pressure (MPa) and temperature (°C), respectively are the most effective parameters on hardness, Young's modulus, shear modulus, apparent porosity and relative density. By comparing the results, a very good agreement was observed between the experimental data and the results obtained from GEP model.

doi*: 10.5829/ije.2020.33.04a.21*

NOMENCLATURE

Wt.% CHA	Weight percent of carbonated hydroxyapatite	R^2	Correlation coefficient
Wt.% Ti	Weight percent of titanium	RMSE	Root mean square error
GEP	gene expression programming	RRSE	Root relative squared error
PSO	particles swarm optimization	MAPE	Mean absolute percentage error
T	Reaction temperature (℃)	N (= 90)	Number of datasets used in the testing and training phases
P	Compaction pressure (MPa)	t_i	The measured values by models
E	Elastic module (Gpa)	p_i	The predicted values by models

1. INTRODUCTION

It is well known that about 65 wt.% of bone is made of Hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$, which is one of the most commonly biocompatible and nontoxic ceramics with the similar chemical and structural characteristics to the human natural bone [1-3]. Therefore, HA is extensively applied for repair and reconstruction of bone tissue defects, making dental, orthopedics and middle ear implants [4, 5], drug delivery and gene delivery [6]. Unfortunately, HA have some disadvantages such as: i. Bone grafting ability of the HA is very slow; ii. The implant is not safe from bacteria; iii. The rate of degradation of this material is slower than the rate of osteogenesis; and iv. The mechanical properties of HA are weaker than living bone [7, 8]. To solve monitored

*Corresponding Author Institutional Email: Khayatireza@gmail.com (G. R. Khayati)

Please cite this article as: M. R. Shojaei, G. R. Khayati, A. Hasani, Utilization of Gene Expression Programming for Modeling of Mechanical Performance of Titanium/Carbonated Hydroxyapatite Nanobiocomposites: The Combination of Artificial Intelligence and Material Science, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 948-955

problems (CO3-2) added to HA. It is clear that except for calcium and phosphorus ions, carbonate ions also in the natural bone mineral, and carbonate ions make up 2 to 8% of the inorganic composition of bone [5]. Therefore, carbonated hydroxyapatite (CHA) has a very similar composition to the human bone. Also, CHA has better solubility and higher biological activity than HA [9-11]. Another way to improve the mechanical properties is the manufacture of nanobiocomposites with some group of metals such as Ti which can be the best choice for the production of CHA nanocomposites due to its unique properties such as relatively low modulus, low density, high strength, corrosion resistance and biocompatibility. However, poor tribological properties, unfavorable mechanical properties, and inability to regenerate bone tissue are the most important disadvantages of Ti [12]. It is noteworthy that CHA and Ti have the ability to compensate for each other's shortcomings and defects, so that combination of these constituents produce a nanocomposite with desirable properties for medical applications. There are various approaches for the preparation of these nanocomposites including sol-gel, co-precipitation, and mechanochemical routes. Among these, the ease of chemical-chemical reduction, high speed, relatively low operating costs, process at the ambient temperature has made it a suitable candidate for the preparation of the composite in this study [13, 14].

Artificial intelligence (AI) based methods such as artificial neural network (ANN), gene expression programming (GEP), and molecular dynamics simulation have been used in many fields, including engineering in recent years [15-17]. For example, GEP has been used extensively in the production of nanocomposites by mechanical alloying method to predict hardness and minimize sintering time [18, 19]. Side by side comparison of literature about the modeling of preparation method are abbreviated in Table 1. To the best of our knowledge, the GEP molding has not been used to optimize the parameters affecting the mechanical properties of Ti/CHA nanocomposites until now. Accordingly, the main contributions of this study are: (a) the usage of GEP to model the consolidation process of preparation of Ti/CHA nanocomposite by mechanochemical approach; (b) assessment of the effect of input parameters such as sintering temperature, Ti and CHA contents on hardness, Young's modulus, apparent porosity, relative density, and theoretical density of the Ti-CHA nanocomposite and (c) the determination and rank of the effect of each practical variable on selected characteristics.

2. OPTIMAZATION APPROACH

This section describes the gene expression programming as a basic concept that is essential to this study.

2. 1. Gene Expression Programming Gene expression programming (GEP), introduced by Ferreira [20] is a new population based evolutionary algorithm that can overcome the disadvantages and limitations of genetic algorithm (GA) and genetic programming (GP) [21]. The GEP encodes the individuals of the created computer programs as linear strings of fixed size (the genome or chromosomes) which are afterwards expressed as nonlinear entities with different sizes and shapes. These entities called as expression trees (ET). Usually these individuals are made up of only one chromosome and each chromosome can have one or more genes. Genes have two main parts: the head and the tail. The head consists of some mathematical operators, variables and constants (*, /, +, -, √, sin, cos, 1, a, b, c) which are used to encode a mathematical expression [18, 19]. The tail just consists of variables and constants (1, a, b, c), which are called terminal symbols. Accordingly, two different languages (Karva Language) are utilized in GEP: the language of the genes and the language of ETs [22]. The translation of Karva to the ET initiates from the leading position in the ET and continues through the string. ET can be translated into the K-expression by registration of the nodes from root layer to the deepest layer [23, 24]. Genes are joined by the linking functions "addition", "subtraction", "multiplication" and "division". For example, an algebraic expression [((a*b)-

TABLE 1. Side by side comparison of literatures about the modeling of preparation of nanocomposite by consolidation of prepared powder by mechanochemical approach

No.	model	Output parameters						Year, [Ref.]
		Compressive strength (MPa)	Porosity (%)	Size of nanoparticle	Hardness(Gpa)	E (Gpa)	Relative density (%)	
1	GEP			✓				[21]
2	PSO			✓				[19]
3	GEP	✓	✓					[16]
4	GEP		✓		✓	✓	✓	This study

c) + √(d-e)] can be shown by a 2 gene chromosome or an expression tree [25]. Figure 1 represents how a chromosome with two genes is encoded as a linear string and how it is expressed as an ET.

3. EXPERIMENTAL METHOD

3. 1. Synthesis and Characterization of Ti-CHA Nanocomposite

Ti powders (Merck, 99 wt.%) with the average particle size 30 mm and CHA were used as mother materials to produce Ti-CHA nanocomposite in various ratios (Table 2) using mechanical alloying. Firstly, the as-received powders of Ti and CHA were mechanically blended in Ar atmosphere using SPEX 8000 for 12 h with the ball-to-powder ratio equal to 1:2 until a good distribution was achieved. The diameter of balls were 10 mm. Secondly, the powders obtained in step 1 were milled for 10 h in a planetary ball mill with rotational speed equal to 400 rpm and ball-to-powder ratio equal to 20:1. Thirdly, the powder was pressed at the pressure of 50 MPa and 80 MPa. After that, the green compacts were sintered at temperatures listed in Table 2.

The apparent porosity of sintered samples was measured by the Archimedes' method using distilled water (ASTM: B962-13) for different temperatures, i.e. 800, 900, 1100 and 1300 °C. To calculate the theoretical density of sintered samples, first relative density was calculated using the measured bulk density (with dense values of Ti and HA as 4.506 and 3.156 g/cm^3, respectively) and then theoretical density was determined for each samples. The Vickers hardness of samples was measured under the load of 10 N for 20 sec. The Lame's constants, i.e. λ and μ were measured by pulse-echo technique MATEC Model MBS8000 DSP (ultrasonic digital signal processing) system with 5 MHz resonating at room temperature. Then, the values of the Young's modulus (E) was calculated from Equations (1)-(3):

$$\lambda = \rho\left(V_L^2 - 2V_S^2\right) \tag{1}$$

$$\mu = \rho V_S^2 \tag{2}$$

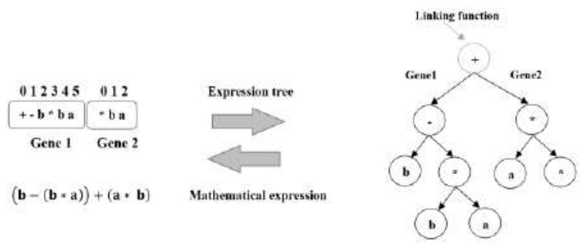


Figure 1. Translating ET language to the mathematical formula [25]

TABLE 2. The details of experimental data for the developed GEP model

Inputs	Variables (Max, Min)
wt.% Ti	24-100
wt.% CHA	0-76
Compaction pressure (MPa)	50-80
Temperature (°C)	800-1200
Output	*
Apparent porosity, %	5.96-8.95
E (GPa)	108.52-143.1
Hardness (GPa)	2.26-2.99
Relative density (%)	88.47-93.23
Theoretical density (g/cm^3)	4.233-4.325

$$E = \mu\frac{3\lambda + 2\mu}{\lambda + \mu} \tag{3}$$

where, V_l^2 and V_s^2 are longitudinal and shear ultrasonic velocities, respectively. *p* is the material bulk density.

3. 2. The Modelling Method

The main parameters of GEP are terminal set, termination condition, fitness function, control parameters and function set [21]. Figure 2 represents the schematic of the GEP algorithm. The process begins with the creation of chromosomes of fixed length for all individuals, randomly. Afterwards, the chromosomes are expressed and the fitness of each individual is investigated. Following to that, the best-fit individuals are chosen to apply the reproduction. The process continues with the new individual for a number of generations until a best solution is found. The genetic operations, e.g. mutation, cross over and reproduction are carried out for the conversion in population [26].

GeneXproTools 5.0 software was used to establish the relationship between input and output parameters, including hardness, elastic modulus, apparent porosity, relative density and theoretical density. In this study, a large number of chromosomes were tested to find the models with the least error. According to the author’s information, GEP has been used to investigate the effect of inputs on several metallurgical outputs**,** which are listed in Table 2. The ranges of parameters involved in the GEP predictive algorithm in this study are summarized in Table 3. To optimize the practical parameters, this study proposes an optimization process using the GEP algorithm.

The performance of the GEP model depends on the number of chromosomes, head sizes, number of genes, linking function, fitness function, mutation, inversion,

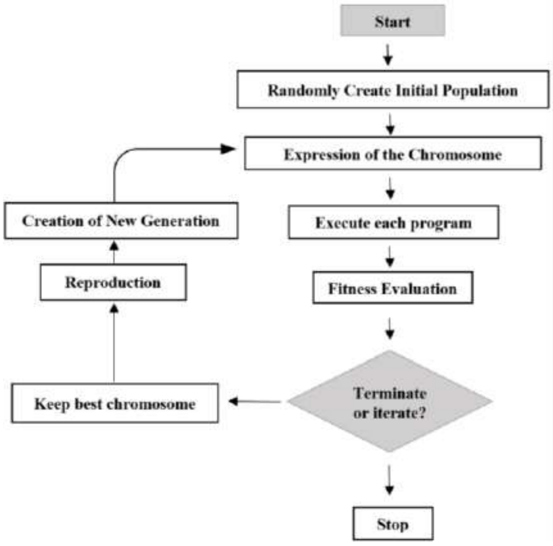


Figure 2. A fundamental flowchart of the GEP algorithm

TABLE 3. GEP parameters settings for the proposed model

GEP parameters definition	Settings
Number of chromosomes	26-30
Head size	8-10
Number of genes	8
Linking function	Multiplication/Division
Fitness function error type	RRSE
Constant per gene	1
Mutation rate	0.0044
Inversion rate	0.15
One point recombination rate	0.2
Two point recombination rate	0.2
Gene recombination rate	0.1
Gene transposition rate	0.1

transposition, constants per gene, number of involving operators, and lower and upper bounds [27]. Table 3 illustrates the features of each model. GEP 1 to GEP 5 models are the most optimal for respectively outputs including hardness, elastic modules, apparent porosity, relative density and theoretical density. The number of functions were changed between 4-17 and in the whole models, the basic operators (+, −, × and /) were presented constantly while the others Sqrt, 3Rt, Exp, Sin, Cos, Atan, Tan, x^2 , x^3 and Ln Csc, Sec, Cot Tanh, Inv, Max^2, Min^2, Avg^2) were added when needed.

4. RESULTS AND DISCUSSION

4. 1. Modelling Observations

To investigate the capabilities of the GEP-based formulation in this study, several statistical parameters were used. Validation of each model by consideration of the mean absolute percentage error (MAPE), root relative square error (RRSE), mean-squared error (MSE) and R square (R^2) were used as the criteria between the experimental and predicted values according to the following Equations (4)-(7).

$$\mathrm{MAPE} = \frac{1}{n}\sum_i \frac{|t_i - p_i|}{t_i} \times 100, \tag{4}$$

$$\mathrm{RRSE} = \sqrt{\frac{\sum_i (t_i - p_i)^2}{\sum_i \left(t_i - (1/n)\sum_i t_i\right)^2}}, \tag{5}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_i (t_i - p_i)^2, \tag{6}$$

$$R^2 = 1 - \left(\frac{\sum_i (t_i - p_i)^2}{\sum_i (p_i)^2}\right), \tag{7}$$

Therefore, t is the experimental (target) value, p the predicted value, and n the data set number in the testing and training phases. If, R^2 values are greater than 0.7 and close to 1 and MAPE, MSE, and RRSE are close to zero, then the results obtained from the models are close to the experimental (target) results [25, 27].

To get generalization capability for the formularization, the experimental data is separated in to two sets as training and test sets. The formularizations are based on training sets and are further tested by test set values to evaluate their generalization capability [28]. All of the results predicted by the training and testing results of GEP-1 to GEP-5 model are given in Table 4. As can be seen, R^2 values for training and testing are in the range of 0.9925-0.9968 and 0.9965-0.9919, respectively, which indicates that the values predicted by GEP are close to the experimental values. The equations obtained for 5 of the best GEP models are summarized in Table 5. Equations are achieved from corresponding expression trees.

The comparison of model predictions against the experimental results of Ti/CHA nanocomposite is shown in Figure 3. It can be seen from Figure 3 that the GEP model could predict the apparent porosity, elastic moduli, hardness, relative density and theoretical density very close to the experimental values. By looking more closely at the graphs, we find that the greatest similarity between the output data and the input data is related to the elastic modulus diagram, which shows that GEP can predict the output parameters of the elastic moduli with the least possible error.

The comparison of model predictions against the experimental results of Ti/CHA nanocomposite is shown in Figure 3. It can be seen from Figure 3 that the GEP

TABLE 3. Characteristics of GEP models

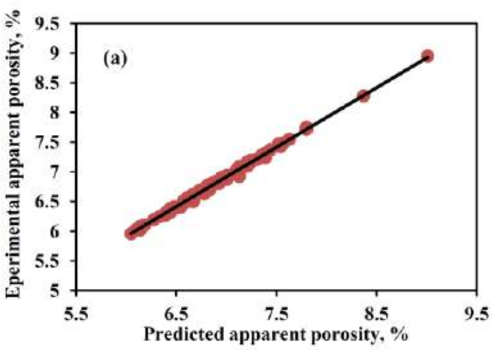
Model	Linking function	Head size	Number of genes	Variable used	Output parameters	Number of functions	Type of function
GEP-1	Multiplication	8	8	Wt.% Ti, Wt.% CHA, P, T	Hardness	6	+, -, *, /, X^2, Exp
GEP-2	Division	10	6	Wt.% Ti, Wt.% CHA, P, T	Elastic modules	17	+, -, *, /, X^2,Ln, Exp, 3RT, Atan, Tanh, Inv, Max^2, Min^2, Avg^2, X^3, X^5 , Sqrt
GEP-3	Division	10	8	Wt.% Ti, Wt.% CHA, P, T	Shear modules	4	+, -, *, /
GEP-4	Multiplication	8	8	Wt.% Ti, Wt.% CHA, P, T	Apparent porosity	7	+, -, *, /, X^2,Ln, Exp
GEP-5	Multiplication	9	7	Wt.% Ti, Wt.% CHA, P, T	Relative density	15	+, -, *, /, X^2,Ln, Exp, Cos, Sin, Tan, Inv, X^3, Csc, Sec, Cot

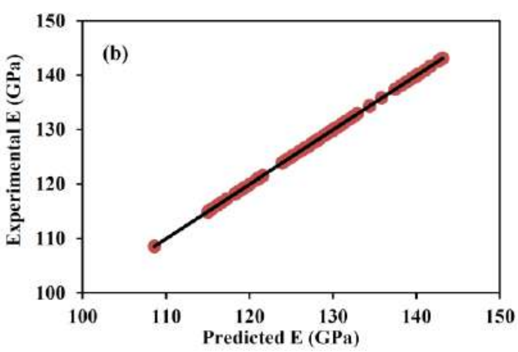
TABLE 4. Statistics of GEP models

No.	R^2		Error					
			Training			Testing		
	Training	Testing	MAPE	MSE	RRSE	MAPE	MSE	RRSE
GEP-1	0.9932	0.9965	5.9	0.0056	0.0913	3.3	0.0022	0.0832
GEP-2	0.9925	0.9941	4.6	0.0035	0.0885	4.2	0.0029	0.0891
GEP-3	0.9968	0.9919	5.08	0.0044	0.1126	4.5	0.0038	0.1022
GEP-4	0.9965	0.9949	5.5	0.0032	0.0929	9.1	0.0119	0.1129
GEP-5	0.9935	0.9932	5.6	0.0041	0.0968	5.8	0.0045	0.1067

TABLE 5. Mathematical equations for GEP model of each parameter

Model	Acquired equation
GEP-1	(9.34/P) * (((0.32-P)+(exp(0.32)/(0.32*P)))^2) * (58.32* T^2)
GEP-2	(((log(0.7)*arctan(%CHA))^2)+exp(log(arctan(%Ti))))*((((%Ti/P)*%CHA)*(%CHA/P))-((0.23*T)*(0.23*T))) * ((((1.0/(%Ti))/(P+2.8))^2)*(arctan(%CHA)-P))
GEP-3	(((T-%Ti)-%Ti)-%CHA)-4*%Ti / (%CHA+(%Ti/(((T+%Ti)*(0.70)-(T-%Ti)))) / ((((%Ti)*(0.53)+(2* %CHA)
GEP-4	((((P+d%Ti)+%Ti)-(2*T))-(exp(-25.66)+((-25.66)-%Ti))) * log((((P*%CHA)+(%Ti+0.27))/((T*0.27)*log(%CHA)))) * ((((-1.6)/(1)/((P^2+(-1.6))-(%Ti+%CHA)));
GEP-5	(0.56) * ((P+sec((0.54*P)))*cos(((1.0/(%Ti))-0/54)) * tan((tan(cos(((reallog(T)+(-0.37))*(-0.37)))) ^{2}))





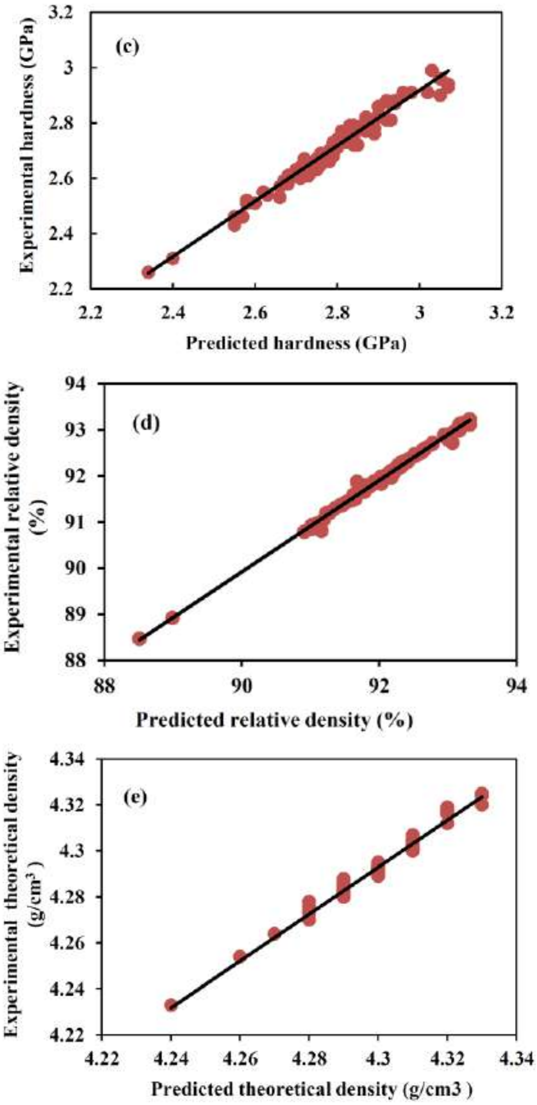


Figure 3. Predicted versus experimental output parameters using GEP model, (a) Apparent porosity, (b) Elastic module, (c) Hardness, (d) Relative density, (e) Theoretical density

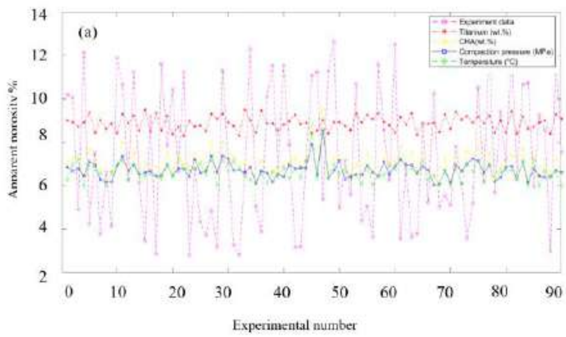
model could predict the apparent porosity, elastic modulus, hardness, relative density and theoretical density very close to the experimental values. By looking more closely at the graphs, we find that the greatest similarity between the output data and the input data is related to the elastic modulus diagram, which shows that GEP can predict the output parameters of the elastic modulus with the least possible error (Table 6).

4. 3. Sensitivity Analysis Finally, sensitivity analysis was used to investigate the effect of input parameters on the output parameters in such a way that the effect of input parameters of output parameter break was performed by keeping the other output parameters constant. Figure 4 (a) shows the effect of input parameters on apparent porosity. As can be seen from Figure 4, the amount of Ti has the greatest effect on the porosity. In other words, porosity increases with decreasing the amount of Ti. In addition, as expected, increasing the pressure reduces the porosity. Figure 4 (b), which examines the effect of input parameters on the elastic modulus, shows that temperature has the least effect on the modulus as compared to other parameters. Examination of experimental data showed that with increasing the amount of CHA, both modulus of elasticity and hardness increase significantly and improve the properties of nanocomposite because Ti alone has good toughness but its hardness is not suitable for medical applications such as fabrication of implants and scaffold. A similar conclusion can be drawn from the graphs related to sensitivity analysis.

In Figures 4 (c, d, e), which are related to the study of the effect of input parameters on the output parameters of hardness, relative density and theoretical density, respectively, it is clear that the amount of titanium has the least effect on these three output parameters compared to other ones. According to the experimental results, it was observed that with increasing the amount of both the relative density and the theoretical density decrease. As mentioned in the previous sections, titanium has a low density so it does not have a significant effect on the density of the nanocomposite.

TABLE 6. Representation of the predicted apparent porosity, elastic, modules, hardness, relative density and theoretical density

Output Dada obtain by GEP model	
Apparent porosity, %	6.05-9.01
E (GPa)	108.6-143.13
Hardness (GPa)	2.34-3.07
Relative density (%)	88.51-93.32



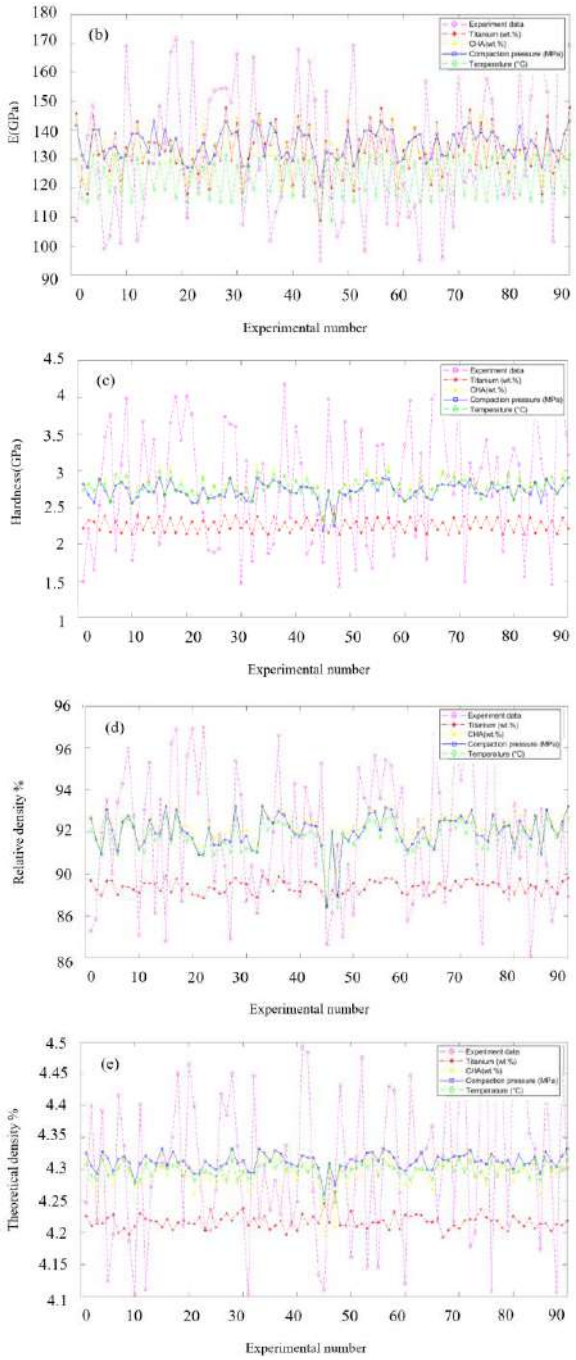


Figure 4. The sensitivity analysis for all parameters, (a) Apparent porosity,(b)Elastic modules, (c) Hardness, (d) Relative density, (e) Theoretical density

5. CONCLUSION

In this research, nanocomposite was produced by mechanical alloying method. After the experimental calculation of apparent porosity, elastic modulus, hardness, relative density and theoretical density as output parameters, GEP modeling was used to estimate the effect of Ti and CHA wt.%, temperature and pressure on nanocomposite properties. Modeling was performed for each of the outputs and the model with the least error was selected for each of them (GEP-1 to GEP-5). For outputs that include porosity, elastic modulus, hardness, relative density, theoretical density. R^2 was obtained for testing (R^2 = 0.9932, 0.9925, 0.9968, 0.9965, 0.9935) and training (R^2 =0.9965, 0.9941, 0.9919, 0.9949, 0.9932), respectively. For these 5 models, R is close to one and the values of MSE and RRSE are close to zero. These results show that Jeep modeling is a very accurate method for predicting the behavior of this nanocomposite. Finally, to be sure, sensitivity analysis was used to investigate the effect of input parameters on each of the output parameters. It was observed that the percentage by weight of Ti has the greatest effect on porosity, sintering temperature has the least effect on the elastic modulus and the percentage by weight of titanium has the least effect on the three outputs of hardness, relative density and theoretical density.

6. REFFERENCES

1. Rahmati, R. and Khodabakhshi, F., "Microstructural evolution and mechanical properties of a friction-stir processed ti-hydroxyapatite (ha) nanocomposite", ***Journal of the Mechanical Behavior of Biomedical Materials***, Vol. 88, No., (2018), 127-139. https://doi.org/10.1016/j.jmbbm.2018.08.025

2. Okamoto, M. and John, B., "Synthetic biopolymer nanocomposites for tissue engineering scaffolds", ***Progress in Polymer Science***, Vol. 38, No. 10-11, (2013), 1487-1503. https://doi.org/10.1016/j.progpolymsci.2013.06.001

3. Takallu, S., Mirzaei, E., Azadi, A., Karimizade, A. and Tavakol, S., "Plate-shape carbonated hydroxyapatite/collagen nanocomposite hydrogel via in situ mineralization of hydroxyapatite concurrent with gelation of collagen at ph= 7.4 and 37° c", ***Journal of Biomedical Materials Research Part B: Applied Biomaterials***, Vol. 107, No. 6, (2019), 1920-1929. https://doi.org/10.1002/jbm.b.34284

4. Bovand, D., Yousefpour, M., Rasouli, S., Bagherifard, S., Bovand, N. and Tamayol, A., "Characterization of ti-ha composite fabricated by mechanical alloying", ***Materials & Design (1980-2015)***, Vol. 65, No., (2015), 447-453. https://doi.org/10.1016/j.matdes.2014.09.021

5. Refaat, A., Youness, R.A., Taha, M.A. and Ibrahim, M., "Effect of zinc oxide on the electronic properties of carbonated hydroxyapatite", ***Journal of Molecular Structure***, Vol. 1147, No., (2017), 148-154. https://doi.org/10.1016/j.molstruc.2017.06.091

6. Prasanna, A. and Venkatasubbu, G.D., "Sustained release of amoxicillin from hydroxyapatite nanocomposite for bone infections", ***Progress in Biomaterials***, Vol. 7, No. 4, (2018), 289-296. https://doi.org/10.1007/s40204-018-0103-4

7. Hannora, A.E. and Ataya, S., "Structure and compression strength of hydroxyapatite/titania nanocomposites formed by high energy ball milling", ***Journal of Alloys and Compounds***, Vol. 658, No., (2016), 222-233. https://doi.org/10.1016/j.jallcom.2015.10.240

8. Cao, Y., Shi, T., Jiao, C., Liang, H., Chen, R., Tian, Z., Zou, A., Yang, Y., Wei, Z. and Wang, C., "Fabrication and properties of zirconia/hydroxyapatite composite scaffold based on digital light

processing", ***Ceramics International***, Vol. 46, No. 2, (2020), 2300-2308. https://doi.org/10.1016/j.ceramint.2019.09.219

9. Kwon, K.-A., Juhasz, J.A., Brooks, R.A. and Best, S.M., "Bioactive conformable hydrogel-carbonated hydroxyapatite nanocomposite coatings on ti-6al-4v substrates", ***Materials Technology***, Vol. 35, No. 11-12, (2020), 727-733. https://doi.org/10.1080/10667857.2018.1475143

10. Youness, R.A., Taha, M.A., El-Kheshen, A.A. and Ibrahim, M., "Influence of the addition of carbonated hydroxyapatite and selenium dioxide on mechanical properties and in vitro bioactivity of borosilicate inert glass", ***Ceramics International***, Vol. 44, No. 17, (2018), 20677-20685. https://doi.org/10.1016/j.ceramint.2018.08.061

11. Ahmed, M., Mansour, S., Al-Wafi, R., Afifi, M. and Uskoković, V., "Gold as a dopant in selenium-containing carbonated hydroxyapatite fillers of nanofibrous ε-polycaprolactone scaffolds for tissue engineering", ***International Journal of Pharmaceutics***, Vol. 577, (2020), 118950. https://doi.org/10.1016/j.ijpharm.2019.118950

12. Niespodziana, K., "Synthesis and properties of porous ti-20 wt.% ha nanocomposites", ***Journal of Materials Engineering and Performance***, Vol. 28, No. 4, (2019), 2245-2255. https://link.springer.com/article/10.1007/s11665-019-03966-8

13. Theingi, M., Tun, K.T. and Aung, N.N., "Preparation, characterization and optical property of lafeo3 nanoparticles via sol-gel combustion method", ***SciMedicine Journal***, Vol. 1, No. 3, (2019), 151-157. https://doi.org/10.28991/SciMedJ-2019-0103-5

14. Gharibshahian, E., "The effect of polyvinyl alcohol concentration on the growth kinetics of ktiopo4 nanoparticles synthesized by the co-precipitation method", ***HighTech and Innovation Journal***, Vol. 1, No. 4, (2020), 187-193. https://doi.org/10.28991/HIJ-2020-01-04-06

15. Gandomi, A.H., Alavi, A.H., Gandomi, M. and Kazemi, S., "Formulation of shear strength of slender rc beams using gene expression programming, part ii: With shear reinforcement", ***Measurement***, Vol. 95, (2017), 367-376. https://doi.org/10.1016/j.measurement.2016.10.024

16. Abdellahi, M., Najafinezhad, A., Ghayour, H., Saber-Samandari, S. and Khandan, A., "Preparing diopside nanoparticle scaffolds via space holder method: Simulation of the compressive strength and porosity", ***Journal of the Mechanical Behavior of Biomedical Materials***, Vol. 72, (2017), 171-181. https://doi.org/10.1016/j.jmbbm.2017.05.004

17. Trang, G.T.T., Linh, N., Linh, N. and Kien, P., "The study of dynamics heterogeneity in sio2 liquid", ***HighTech and Innovation Journal***, Vol. 1, No. 1, (2020), 1-7. https://doi.org/10.28991/HIJ-2020-01-01-01

18. Ebrahimi-Kahrizsangi, R., Abdellahi, M. and Bahmanpour, M., "Ignition time of nanopowders during milling: A novel simulation", ***Powder Technology***, Vol. 272, (2015), 224-234. https://doi.org/10.1016/j.powtec.2014.12.009

19. Khayati, G.R., "A predictive model on size of silver nanoparticles prepared by green synthesis method using hybrid artificial neural network-particle swarm optimization algorithm", ***Measurement***, Vol. 151, (2020), 107199. https://doi.org/10.1016/j.measurement.2019.107199

20. Ferreira, C., "Gene expression programming: A new adaptive algorithm for solving problems", arXiv preprint cs/0102027, (2001). https://arxiv.org/abs/cs/0102027

21. Sattari, R. and Khayati, G.R., "Prediction of the size of silver nanoparticles prepared via green synthesis: A gene expression programming approach", ***Scientia Iranica***, Vol. 27, No. 6, (2020), 3399-3411. https://dx.doi.org/10.24200/sci.2020.53209.3112

22. Ghayour, H., Abdellahi, M. and Bahmanpour, M., "Artificial intelligence and ceramic tools: Experimental study, modeling and optimizing", ***Ceramics International***, Vol. 41, No. 10, (2015), 13470-13479. https://doi.org/10.1016/j.ceramint.2015.07.1387

23. Abdellahi, M., Bahmanpour, H. and Bahmanpour, M., "The best conditions for minimizing the synthesis time of nanocomposites during high energy ball milling: Modeling and optimizing", ***Ceramics International***, Vol. 40, No. 7, (2014), 9675-9692. https://doi.org/10.1016/j.ceramint.2014.02.049

24. Sabouhi, R., Ghayour, H., Abdellahi, M. and Bahmanpour, M., "Measuring the mechanical properties of polymer–carbon nanotube composites by artificial intelligence", ***International Journal of Damage Mechanics***, Vol. 25, No. 4, (2016), 538-556. https://doi.org/10.1177%2F1056789515604375

25. Ebrahimi-Kahrizsangi, R., Abdellahi, M. and Bahmanpour, M., "Self-ignited synthesis of nanocomposite powders induced by spex mills; modeling and optimizing", ***Ceramics International***, Vol. 41, No. 2, (2015), 3137-3151. https://doi.org/10.1016/j.ceramint.2014.10.158

26. Gholampour, A., Gandomi, A.H. and Ozbakkaloglu, T., "New formulations for mechanical properties of recycled aggregate concrete using gene expression programming", ***Construction and Building Materials***, Vol. 130, (2017), 122-145. https://doi.org/10.1016/j.conbuildmat.2016.10.114

27. Jafari, M.M. and Khayati, G.R., "Prediction of hydroxyapatite crystallite size prepared by sol–gel route: Gene expression programming approach", ***Journal of Sol-Gel Science and Technology***, Vol. 86, No. 1, (2018), 112-125. https://link.springer.com/article/10.1007/s10971-018-4601-6

28. Kök, M., Kanca, E. and Eyercioğlu, Ö., "Prediction of surface roughness in abrasive waterjet machining of particle reinforced mmcs using genetic expression programming", ***The International Journal of Advanced Manufacturing Technology***, Vol. 55, No. 9, (2011), 955-968. https://link.springer.com/article/10.1007%252Fs00170-010-3122-4

Persian Abstract

چکیده

نانوبیوکامپوزیت‌های هیدروکسی آپاتیت کربنات- تیتانیوم به دلیل زیست‌سازگاری فوق‌العاده و خصوصیات مشابه آپاتیت استخوان، کاربردهای بیولوژیکی بسیاری مانند بیو ایمپلنت دارند. کربنات هیدروکسی آپاتیت از ویژگی‌های بیولوژیکی خوبی برخوردار است اما بزرگ‌ترین مشکل نانوکامپوزیت‌های Ti-CHA ویژگی‌های مکانیکی آن‌ها مانند سختی، مدول یانگ، تخلخل ظاهری و تراکم نسبی است. پژوهش حاضر، با استفاده از برنامه نویسی بیان ژن (GEP)یک مدل پیش بینی می¬شود. از GEP برای بهینه سازی تأثیر فشار پرس، CHA و Ti و دمای تف‌جوشی بر خواص مکانیکی کنترل شده آن‌ها استفاده می‌شود. برای دستیابی به این هدف 90 آزمایش مختلف برای ایجاد مدل‌های GEP در نظر گرفته شد. پس از آن نتایج به طور تصادفی به 63 مجموعه آموزش و 27 مجموعه آزمون تقسیم شدند. سرانجام 5 مورد از بهترین مدل‌ها برای هر پارامتر خروجی متفاوت گزارش شدند. آنالیز حساسیت برای بررسی تاثیر پارامترهای عملی بر روی هر یک از خصوصیات کامپوزیت مانند سختی، مدول یانگ، تخلخل ظاهری و چگالی نسبی انجام شد. با مقایسه نتایج مطابقت قابل قبولی بین پارامترهای خروجی تجربی و پارامترهای به دست آمده از مدل GEP و آنالیز حساسیت مشاهده شد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 956-965

International Journal of Engineering

Journal Homepage: www.ije.ir

Rejuvenation Heat Treatment of Nickel Base Superalloy Grade GTD111 after Long-term Service via Taguchi Method for Optimization

S. Toocharoen, S. Kaewkuekool, P. Peasura*

Department of Production Technology Education, Faculty of Industrial Education and Technology, King Mongkut's University of Technology Thonburi, Thailand

P A P E R I N F O

Paper history:
Received 18 December 2020
Received in revised form 27 January 2021
Accepted 02 February 2021

Keywords:
GTD111
Rejuvenation Heat Treatment
Taguchi Method
Hardness
Microstructure

A B S T R A C T

This research describes an optimization and rejuvenation of the heat treatment process for a nickel base superalloy grade GTD111 after long-term service. The aging heat treatment variables examined in this study included primary aging temperature, primary aging time, secondary aging temperature, and secondary aging time. The resulting materials were examined using Taguchi method design of experiments to determine the resulting material hardness test and observed with the hot tensile test, scanning electron microscopy, and energy dispersive X-ray spectroscopy. The experimental results showed what happens following optimization with the heat treatment parameters of a primary aging temperature of 1120 °C, primary aging time of 3 h, secondary aging temperature of 845 °C, and secondary aging time of 24 h. The material, after rejuvenation heat treatment via optimization with γ′ particle characteristics, had a coarse square shape, spherical shape of γ′, and fine γ′ precipitate distributed on the parent phase, which affects the mechanical properties of the material. fine γ′ precipitate distributed on parent phase, which affects the mechanical properties of the material.

***doi**: 10.5829/ije.2021.34.04a.22*

1. INTRODUCTION

Nickle base superalloys grade GTD111 are used extensively in the industry for high-temperature applications such as the production of the gas turbine. The GTD111 grade is used for high-temperature applications because it has good mechanical properties and is resistant to corrosion and fatigue at high temperatures. Esanab et al. [1] conducted a data-based investigation on the performance of an independent gas turbine and found that turbines are one of the most prominent technologies being adopted in producing electricity from natural gas; this part is attacked by hot gas. The mechanical properties of superalloys create a multiphase, such as precipitation strengthening of $Ni_3(Al,Ti)$ or γ' phase and carbide phase. The particles disrupt the movement of the dislocation; this increases the strength of the material [2,3].

The microstructural characteristics of the alloy would be normally degraded, resulting in poorer mechanical properties after long-term service and finally leading to failure of components. These increasingly extreme operating environments accelerate the aging process in materials, leading to reduce performance and hot corrosion failure. The cause of blade trailing edge failure is thermal stress leading to thermal fatigue according to Perrut et al. [4] and Salehnasaba et al. [5]. Zhangab et al. [6] studied creep behavior and deformation mechanisms of a novel directionally solidified Ni-base superalloy and found that to repair the material, it is necessary to restore its initial properties and the original microstructure first. The alloy gains its appropriate microstructure and high-temperature strength through a precipitation hardening mechanism. Heat treatment is the way to improve this material.. Kim et al. [7] studied transient liquid phase bonding of γ′- precipitation strengthened Ni-based superalloys. Rezaie et al. [8] studied the effects of temperature and pressure in the HIP process on mechanical properties and found that homogenization treatment resulted in a more uniform distribution of the

*Corresponding Author Email: *prachya.pea@kmutt.ac.th* (P. Peasura)

Please cite this article as: S. Toocharoen, S. Kaewkuekool, P. Peasura, Rejuvenation Heat Treatment of Nickel Base Superalloy Grade GTD111 after Long-term Service via Taguchi Method for Optimization, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 956-965

γ′ precipitates. This shows that precipitation hardening of γ′ precipitates affects the microstructure. In their studies, the heat treatment conditions influenced the size distribution and the microstructure of γ′ precipitate, which mainly include solution heat treatment and aging. All the parameters which are important for repair, such as time and the temperature of the heat treatment process, affected the microstructure. In a study on the effect of reheat treatment on microstructure and stress rupture property of a directionally solidified material, Jiang et al. [9] studied temperature and pressure in the HIP process on the mechanical properties of GTD-111 after long-term thermal exposure. Kaplanskii et al. [10] reported an influence of aging and HIP treatment on the structure and properties of NiAl-based turbine blades manufactured by laser powder bed fusion, and Xua et al. [11] studied short-term creep behavior of an additive manufactured non-weldable Ni-base superalloy evaluated by slow strain rate testing.

The method for improving the mechanical properties and microstructure of the GTD111 grade is extremely important. Statistical studying is not used only for data analysis, but primarily for the planning of experiments. Taguchi's method can determine the best combination of factors and interact with the influence behavior of the response variable in the given process. C.D. Pimenta et al. [12] studied the application of Taguchi's method in the investigation of influential heat treatment factors in steel wires. Taguchi's method can be used to create mathematical equations for predicting the mechanical properties of heat treatment. H. Terzioglu [13] analyzed the effect factors of Taguchi's method and its application for the optimization of Ni-base superalloys. G. kartheeka et al. [14] and A. Kumar et al. [15] studied the optimization of residual stresses in the heat treatment process of superalloys. R.M. Dodoa et al. studied multi-response optimization and analysis through Taguchi's method for quenching applications [16]. M. Sobhani et al. [17] and A.S. Canbolata et al. [18] used Taguchi's method for optimization to find the best parameter of experiment. Despite all these studies, no research has been applied to the Taguchi experiment to identify the efficiency of the heat treatment process in improving the mechanical and microstructural properties of GTD 111 after long-term service.

In this study, Ni-base superalloys grade GTD-111 which had been degraded following a long-term service period were used. The heat treatment process was used for rejuvenation in nickel base superalloys grade GTD111. Therefore, the research studied the rejuvenation of materials with heat treatment after long-term use in order for them to be recycled, as well as the effects of applying our experimental design combining Taguchi's design method and the optimization and mathematic model on the mechanical properties and microstructure refurbishment of grade GTD111 materials. The study also examined the effects of heat treatment factors on the GTD111 grade performance characteristics, such as microstructure, hardness, and strength, with the hot tensile test. It is also essential to understand the process behavior followed by a proper selection of heat treatment factors of critical importance to achieve satisfactory results in terms of extending the useful life of the material, as well as a significant cost reduction, and be able to apply the material in the gas turbine engine industry with the highest quality and efficiency.

2. MATERIALS

2. 1. Material GTD111 after long-term service was used for the test specimen shown in Figure 1. The specimens were 5.00 mm thick. The details of the chemical composition of the GTD111. GTD111 are provided in Table 1.

2. 2. Rejuvenation Heat Treatment Process This research examined the rejuvenation of used GTD111 material via the heat treatment process. The heat treatment process used to improve the microstructure and mechanical properties has 3 steps: first, solution heat treatment, second, primary aging, and third, secondary aging. For solution aging the research was controlled as 1200 °C at 4 hours for every sample. The research studied two factors, the influence of primary aging and secondary aging, because both factors affect the γ ' size and the mechanical properties of the material. The sequence of the heat treatment process can be seen in Figure 2.

2. 3. Design of Experiment The experimental design through Taguchi's method is the application of the experimental design to control factors, uncontrollable

Figure 1. Nickel base superalloys grade GTD111 after long-term service

TABLE 1. Chemical composition of the GTD111 by weight (%)

Ni	Cr	Co	Ti	W	Al	Ta	Mo	Fe	C	B
Bal.	13.5	9.5	4.75	3.8	3.3	2.7	1.53	0.23	0.09	0.01

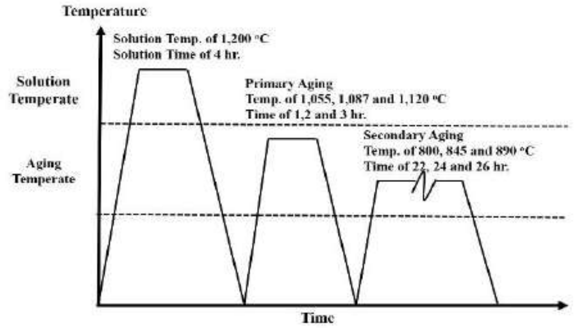


Figure 2. The sequence of the rejuvenation heat treatment process

factors or noise factors, all of which are variables that serve as the source of variation. It is not possible to eliminate the influence of these variables. Therefore, the main function of Taguchi's method was to reduce data fluctuations by selecting the control factors. The raw experiment results were converted into a signal to noise (S/N) ratio which is important in finding the right target. To optimize the characteristics of the S/N ratio, it can be divided into 3 types, which are "Small the Better Type Problem", "Nominal the Best Type Problem", and "Larger the Better Type Problem". [18] The benefits of the Taguchi method are that it helps to reduce the number of trials, saves time, reduces trial costs and gives reliable results. The research procedure for the Taguchi's method experiment is shown in Figure 3.

For the design of the experiment with the Taguchi method, we considered the factors that affect the hardness in the heat treatment process using the L27 (34) orthogonal array. There are 4 input factors, which are primary aging temperature, primary aging time, secondary aging temperature, and secondary aging time, each of which is divided into 3 levels of experiments. A treatment total of 27 is shown in Table 2.

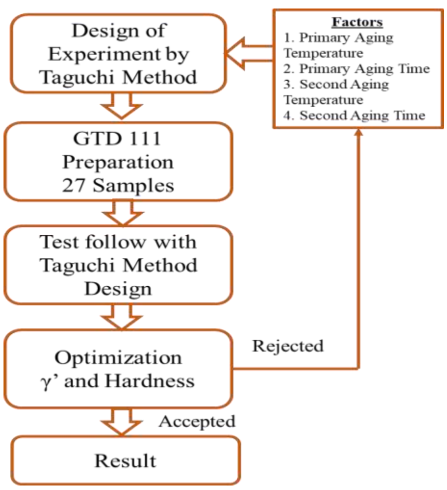


Figure 3. The research procedure of the Taguchi method

TABLE 2. Experimental Taguchi Design

Factors	Symbol	Level		
Primary Aging Temperature	X_1	1.055	1.087	1.120
Primary Aging Time	X_2	1	2	3
Secondary Aging Temperature	X_3	800	845	890
Secondary Aging Time	X_4	22	24	26

2. 4. Hardness Test This research used the Vickers hardness test method. Vickers hardness consists of indenting the test material with a diamond pyramid with a square base and an angle of 136 degrees. The two diagonals of the indentation left on the surface of the material after removal of the load are measured using a microscope, and their average is calculated. The area of the sloping surface of the indentation is calculated. Vickers hardness is the quotient obtained by dividing the kg_f load by the square mm area of indentation. Hardness was analyzed with the Vickers hardness tester Model MMT-X7 and was determined using a 500 g_f for dwell time of 10 sec to observe the effects of the heat treatment factors.

2. 5. Microstructure Test The specimens were sectioned transversely and polished using standard metallographic techniques. The microstructures were examined and analyzed with an optical microscope (OM). Additionally, scanning electron microscopy (SEM) also used to examine the specimens. Etching was performed using Marble's reagent (7.5 mL HF, 2.5 mL HNO3, 200 mL methanol) for 30 seconds to permit observation of the microstructure with scanning electron microscopy. The microstructures in the γ' size were examined and analyzed with an ImageJ program.

2. 6. Hot Tensile Test The heat treatment samples were hot tensile tested to analyze the mechanical properties of Ni-Base super alloys grade GTD111. The tensile fracture characteristics of the samples with the best mechanical properties of hardness and weakness and mechanical properties of hardness were observed to confirm the critical temperature of high-temperature embrittlement fracture. The specimens were tensile tested at a high temperature of 871 °C [19] with the universal testing machine model HT-8336. The microstructure of the hot tensile specimens was observed by scanning electron microscopy. The dimensions of the hot tensile samples are shown in Figure 4.

3. RESULT AND DISCUSSION

3. 1. Taguchi Method Analysis The Taguchi method was designed form orthogonal array experiments

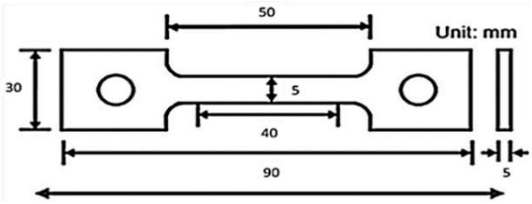


Figure 4. Dimensions of a hot tensile samples

of 27 experiments for 4 factors and 3 levels, considering the factors that affected the average of the interested response, the hardness, and the average signal to noise (S/N) ratio. [17,18] The research divided the Taguchi problem functions into categories, namely, the larger the better type for hardness as calculated from Equation (1). The experimental results showed the hardness and average signal to noise ratio are shown in Table 3.
Problem larger is better type:

$$S/N_L = 10log\left[\frac{1}{n}\sum\frac{1}{y_i^2}\right] \quad (1)$$

where Y_i is the data measured in duplication. n is the repeated trial number.

The results showed that primary aging temperature (X_1) had the highest affected on response for signal to noise ratios in rejuvenation heat treatment process which followed by secondary aging temperature (X_3), primary aging time (X_2) and secondary aging time(X_4), respectively (Table 4). In these experiments, the difference between the average value of each factor is considered a level of the factor.

TABLE 3. Experiment result for hardness based on Taguchi's L27 orthogonal array design matrix

Order	X_1	X_2	X_3	X_4	Hardness (HV)	S/N
1	1055	1	800	22	496.40	53.91
2	1055	1	845	24	501.40	54.12
3	1055	1	890	26	499.60	53.97
4	1055	2	800	24	502.90	54.03
5	1055	2	845	26	506.00	54.08
6	1055	2	890	22	498.40	53.95
7	1055	3	800	26	505.50	54.07
8	1055	3	845	22	510.90	54.16
9	1055	3	890	24	501.40	54.00
10	1087	1	800	22	502.90	53.96
11	1087	1	845	24	511.90	54.18
12	1087	1	890	26	502.40	54.02
13	1087	2	800	24	512.30	54.19
14	1087	2	845	26	512.60	54.19
15	1087	2	890	22	507.90	54.11
16	1087	3	800	26	513.30	54.20
17	1087	3	845	22	518.00	54.28
18	1087	3	890	24	514.00	54.21
19	1120	1	800	22	516.50	54.26
20	1120	1	845	24	519.10	54.30
21	1120	1	890	26	515.30	54.24
22	1120	2	800	24	529.50	54.37
23	1120	2	845	26	524.80	54.40
24	1120	2	890	22	512.40	54.19
25	1120	3	800	26	526.80	54.43
26	1120	3	845	22	528.10	54.45
27	1120	3	890	24	517.40	54.22

TABLE 4. Response for signal to noise ratios

Level	X_1	X_2	X_3	X_4
1	54.04	54.11	54.16	54.15
2	54.15	54.17	54.24	54.18
3	54.32	54.23	54.10	54.18
Delta	0.29	0.12	0.14	0.04
Rank	1	3	2	4

The results of the variance analysis to test the hypothesis in the research determined the confidence level at 95 percent (p-value <0.05). The analysis of the experimental results is shown in Table 5. The research found that primary aging temperature, primary aging time, and secondary aging temperature have p-values ≤ 0.001, 0.001, and 0.001, respectively, indicating that the 3 factors have a significant (p-value <0.05) influence on hardness in the heat treatment process.

From Table 6, the coefficient for hardness of the heat treatment of GTD111 shows that primary temperature and secondary temperature have an effect on hardness at 95% confidence level (p-value <0.05). The heat treatment factors that did not affect hardness were secondary temperature and secondary time. The results indicate that the heat treatment data could be hardness predicted using the model shown in model 2. The R^2(adj) of the collected data was approximately 83.72%, which shows that the response can be described by the experimental factors.

The regression model for hardness is:

$$Y_{HV} = 220 + 0.279(X^1) + 3.880(X^2) - 0.0362\,(X^3) + 0.414\,(X^4) \quad (2)$$

3. 2. Optimization for GTD111 Heat Treatment Process

The determination of the optimization of 4 factors in the heat treatment process is shown in Figure 5, which depicts the relationship of the averages and the response to the factors. The research selected the S/N ratios of the larger the better type for hardness.

TABLE 5. Analysis of variance for signal to noise (SN) ratios of hardness

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Primary Aging Temp	2	0.370	0.370	0.185	62.640	0.001
Primary Aging Time	2	0.065	0.065	0.032	11.070	0.001
Secondary Aging Temp	2	0.088	0.088	0.044	15.000	0.001
Secondary Aging Time	2	0.008	0.008	0.004	1.450	0.261
Residual Error	18	0.053	0.053	0.002		
Total	26	0.586				

TABLE 6. Estimated model coefficients for hardness

Term	Coef	SE Coef	T	P
Constant	220.150	34.520	6.380	0.001
Primary Temp (X_1)	0.279	0.026	10.590	0.001
Primary Time (X_2)	3.883	0.857	4.530	0.001
Secondary Temp (X_3)	-0.036	0.019	-1.900	0.071
Secondary Time (X_4)	0.413	0.428	0.970	0.345
$R^2 = 86.20\%$		$R^2_{(adj)} = 83.72\%$		

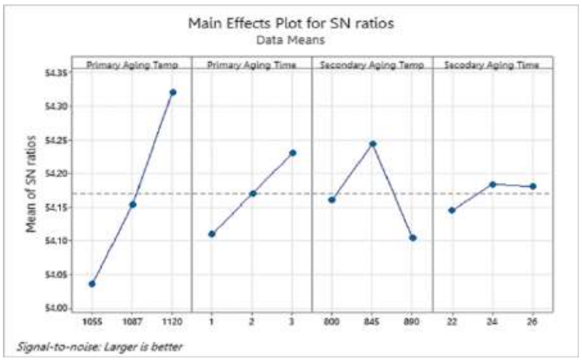


Figure 5. Main effects plot for S/N ratios of hardness

For the variation between the factors, the signal to noise ratio (S/N) was calculated following the methodology. The relationship “larger is better” was adopted in this study. The result showed that he highest hardness value was observed at 1120 °C for 3 h in primary aging temperature and primary aging time, respectively. However, the hardness value was decreased with increasing temperature from 845-890 °C which the highest hardness value was observed at 845 °C in secondary aging temperature. The secondary aging time at 24 h. had the highest hardness value which hardness decreased after 24 h. It indicated the change in hardness from graph as a result of over aging.[3-4] The aging heat treatment factors have the highest hardness, shown in Figure 5, with a primary aging temperature of 1120 °C, primary aging time of 3 h, secondary aging temperature of 845 °C, and secondary aging time of 24 h.

Taguchi’s method was analyzed, and the optimization found that a primary aging temperature of 1,120 °C, primary aging time of 3 h, secondary aging temperature of 845 °Cm and secondary aging time of 24 h result in the maximum GTD111 hardness. Figure 6A depicts the microstructures of the GTD111 grade specimen after long-term service at expanding 50,000 h. For the

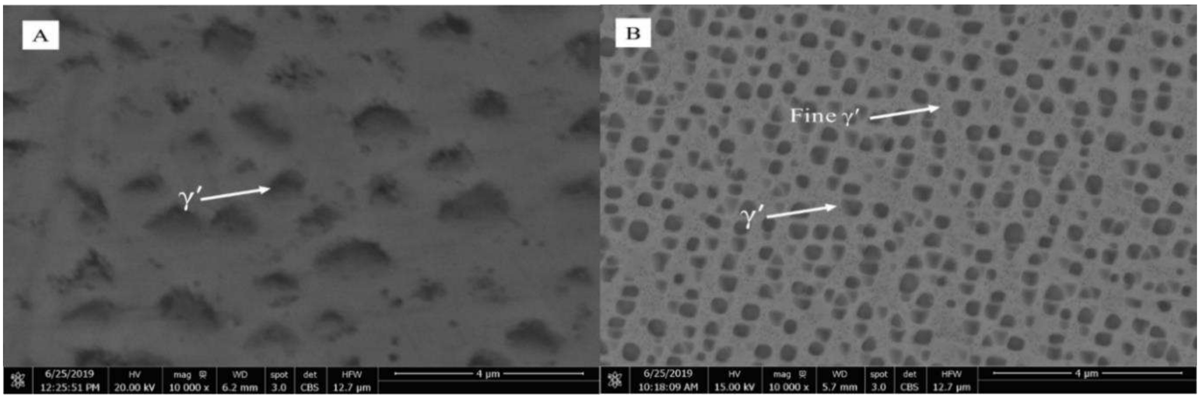


Figure 6. SEM image, showing the GTD111 grade. A. after long-term service at expanding 50,000 h. B. primary aging temperature of 1,120 °C, primary aging time of 3 h, secondary aging temperature of 845 °C, and secondary aging time of 24 h

microstructure, an austenite (γ) phase matrix comprising the main structural and scattered γ′ asymmetrical shape was observed in the parent phase specimen with a hardness of 430.30 HV. The optimization condition in Figure 6B shows that the square shapes of γ′ and fine γ′ precipitates were distributed on the γ phase, which had a γ′ size of 0.38 µm and hardness of 530.10. A comparison of the microstructures of the material after long-term service and the aging heat treatment specimens revealed clear differences. There was a reform γ′ shape and size increase in density of the square shape γ′ and fine γ′, which resulted in increased GTD111 hardness consistent with the research. Rezaiea [20] also reported that the size shape of γ′ precipitates and mechanical property was controlled by aging temperature and time. Therefore, the optimization condition was also used in the improvement of the microstructure and hardness of GTD111.

3. 3. Microstructure Analysis The microstructure was examined via a scanning electron microscope (SEM) and energy dispersive X-ray spectrometer (EDS) of the GTD111. The microstructure showed coarse and fine γ' particle phases with various shapes, including a spherical shape, square shape, and cubic shape of γ' particles.

.

In Figure 7, where an SEM image shows the microstructure of superalloys grade GTD111 after aging heat treatment, we can see the gamma prime particle (γ′ or Ni3(Al,Ti))[21] precipitate evenly distributed in the face centered cubic (FCC) or gamma (γ) matrix [22]. Analyzing the elemental concentrations in the tertiary γ′, EDS as shown in Figure 8 exhibits additional related elements such as Cr Co, containing a large number of elements. The γ phase was decorated with a discontinuous carbide layer scattered throughout the matrix. The carbides were MC (Ta,Ti,W) [21]. The MC carbides were evenly distributed throughout the matrix [21, 9].

The nickel base superalloys were tensile tested, hardness tested, creep tested, etc. to evaluate their mechanical properties. However, as the temperatures increased, the differences between the tensile properties of different alloys almost disappeared. For polycrystalline nickel-based superalloys, there are mainly three strengthening mechanisms, which are γ′ precipitate strengthening, γ phase solid solution strengthening, and grain boundary strengthening form carbide. The primary aging temperature of 1120 °C for 3 h and secondary aging temperature of 845 °C for 24 h. Figure 9 A shows the microstructure where the carbide,

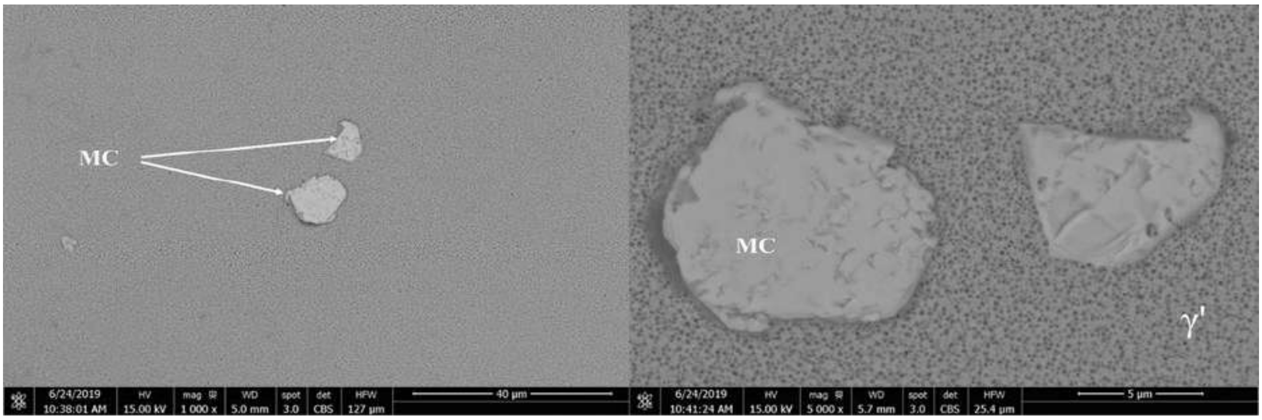


Figure 7. SEM image—the microstructure of GTD111 after the aging heat treatment process

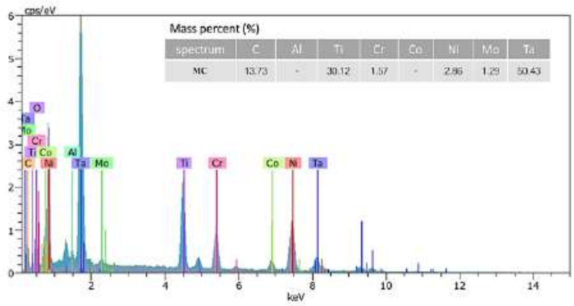


Figure 8. Quantitative analysis of particle composition by EDS

which is scattered along the line, continues to have a carbide structure on the γ′ and γ phases. The chemical composition of the primary MC carbides in the heat-treated alloy measured is illustrated in Figure 8. The main elements forming the primary MC carbides are Ti and Ta; the MC carbides show little morphological change after reheat treatment. Figure 9B shows the primary aging temperature of 1120 °C for 3 h and secondary aging temperature at 890 °C for 24 h. The SEM scattered image of the carbide structure with γ′ precipitate with a different size. The dispersion of these constituents formed from non-equilibrium solidification can degrade the tensile

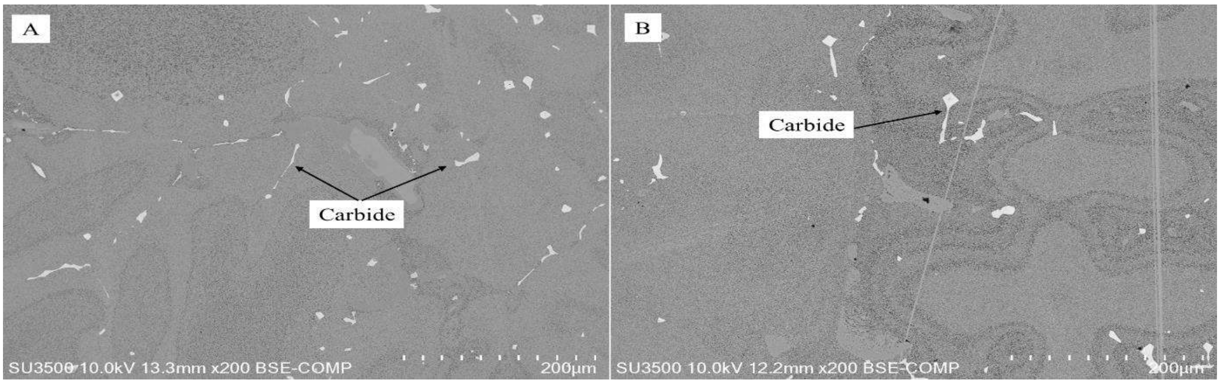


Figure 9. SEM of GTD111 grade after the heat treatment process. A. primary aging temperature of 1,120 °C for 3 h., secondary aging temperature of 845 °C for 24 h. B. Primary aging temperature of 1,120 °C for 3 h., secondary aging temperature of 890 °C for 24 h

properties at high temperatures. The reheat-treated alloy was dangerous as it risked causing intergranular fracture along the grain boundaries, where the γ′ precipitates are rafted and in the interdendritic region where MC carbides are located. However, cracking can form at the interdendritic region; some cracks also grew in the γ′ phase, which indicated bigger grain boundary strengths than reported in the corresponding results of Long [23] and Berahmand [24]. The different parameters after the heat treatment process of Ni-base superalloy grade GTD 111 can influence variations in the microstructure and mechanical properties.

The aim of the experiment was to improve the hardness and microstructure with the heat treatment process according to the 27 samples of the Taguchi experiment design, then test the microstructure with SEM and EDS. Low heat input in primary heat treatment with a primary aging temperature of 1050 °C for 1 h. and secondary aging temperature of 800 °C for 22 h. found that the small spherical and square shapes of the γ′ particle were small, with a size of 0.086 µm, and the γ′ particle precipitate was evenly distributed on the γ phase, as shown in Figure 10A. The samples tested with a primary aging temperature of 1087 °C for 2 h. and secondary aging temperature of 845 °C for 26 h. found that for the γ′ size of 0.270 µm, the particle characteristics were a coarse square shape of γ′ and an evenly distributed precipitate, as shown in Figure 10B. In the SEM microstructure in Figure 10C, with a primary aging temperature of 1120 °C for 3 h. and secondary aging temperature of 845 °C for 24 h. A coarse square shape, spherical shape of γ′ and fine γ′ precipitate distributed on the parent phase was found, in agreement with previous studies of Zhang [25]. At a secondary aging increase of 890 ° C, where the coarse γ ′ of square shape increased in size, there was a fine distribution of γ ′ in the less dense parent phase, as shown in Figue 10D. The result of over-aging temperature causing the lowest hardness of all of the GTD111 is in accordance with the research of Wang [25] and Zhanga [26], who reported that the increase in the secondary aging temperature results in an increase on the coarse γ′ size, which affects the mechanical properties of the material.

3. 4. Hot Tensile Analysis The material GTD111 was used at temperatures between 750 and 900 °C in the combustion chamber of the jet engine [27]. According to the research, once the mechanical properties and microstructural refurbishment of the GTD111 material were modified with rejuvenation heat treatment, it was necessary to conduct tests to confirm the feasibility of using heat-treated materials with hot tensile strength tests. The hot tensile strength test was selected for the optimum conditions from the Taguchi experiment at a primary aging temperature of 1120 ° C for 3 h. and secondary aging temperature of 845 ° C for 24 h. The specimen with the lowest hardness at a primary aging temperature of 1050 ° C for 1 h. and secondary aging temperature of 800 ° C for 22 h. was tested at 871° C, in accordance with ASTM E8.

The hot tensile test showed that for the GTD111 grade, in an optimization condition of primary aging temperature of 1120 °C for 3 h. and secondary aging temperature sof 845 °C for 24 h. The specimen had a higher strength at 402 MPa. A comparison was run with a GTD111 heat treatment with a primary aging temperature of 1050 ° C for 1 h and secondary aging temperature of 800 °C for 22 h, and the tensile strength was found to be 316 MPa. Figure 11A shows the fracture mechanism, which shows rather brittle characteristics, and GTD111 also has a continued cleavage fracture behavior in the microcrack area. This facture mode was associated mainly with the secondary phase. Dadkhah and Kermanpur [28] also reported corresponding results, such as that the primary aging temperature changed the fracture mode from ductile to rather brittle and also

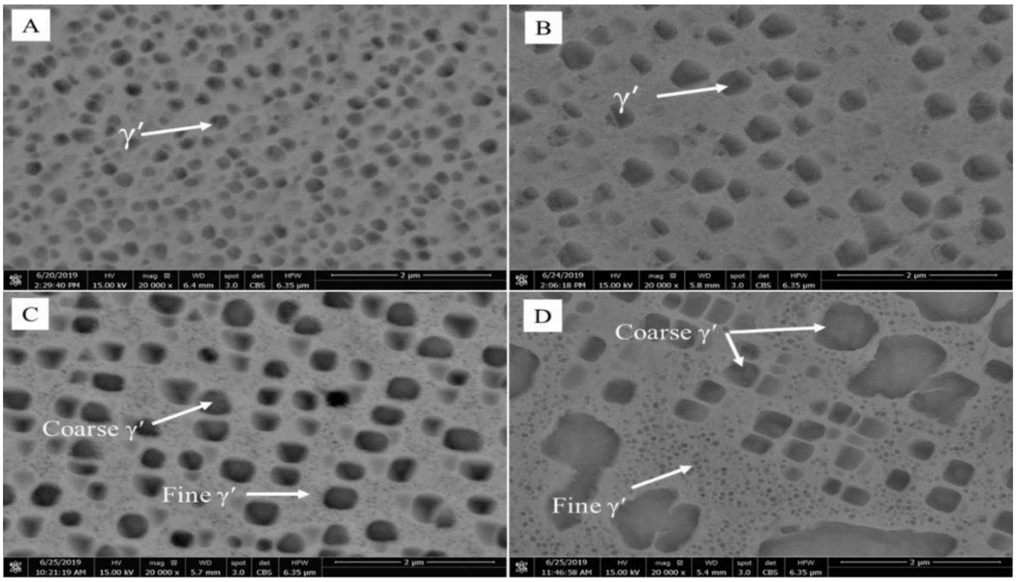


Figure 10. SEM image—the microstructure of GTD111 after heat treatment process. A. Primary aging temperature of 1,050 °C for 1 hr, secondary aging temperature of 800 °C for 22 h. B. Primary aging temperature 1,087 °C for 2 h., secondary aging temperature of 845 °C for 26 h. C. Primary aging temperature of 1,120 °C for 3 h., secondary aging temperature of 845 °C for 24 h. D. Primary aging temperature of 1,120 °C for 3 h., secondary aging of 890 °C for 24 h

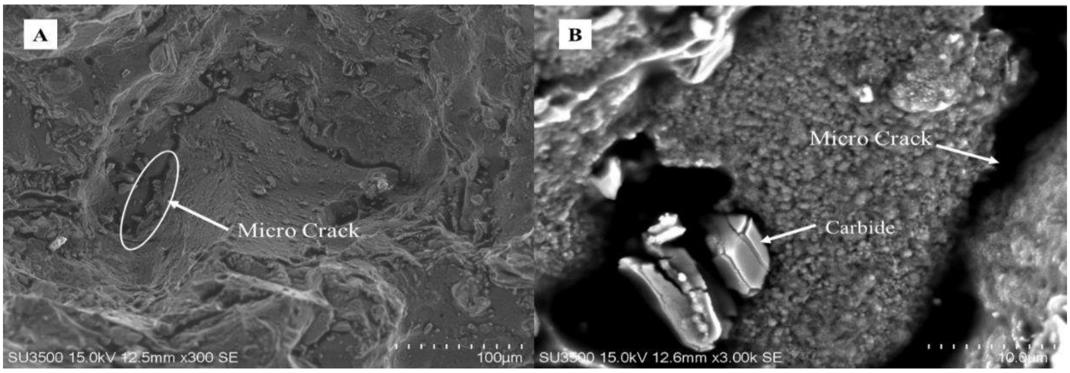


Figure 11. Fracture surface of GTD111 after a hot tensile test of 871°C

consistent with the higher hardness in the specimen. This is also consistent with the findings of Sajjadi et al. [29], who studied tensile deformation mechanisms at different temperatures in the Ni-base superalloy GTD-111, which resulted in hot tensile deformation and led to growth of γ′ particles normal and carbide in the load direction. Figure 11 (B) shows a zoomed image of the fracture area, where it was found that the MC carbides are near the micro crack. The negative lattice misfit of γ/γ′ and the position of the carbide had an effect on the loading direction in the hot tensile test [30]. The MC carbide in the thermally exposed alloy resulted in cracks being formed along the grain boundaries, which is similar to the corresponding results of. Choi et al. [31]. The behavior in carbide phases resulted in fractures after the hot tensile test because primary MC carbides were scattered within the grain interiors as well as along the grain boundaries.

4. CONCLUSION

This research work examines the effects of heat treatment on grade GTD111 after long-term service using optimization with Taguchi method. Primary aging temperature, primary aging time, secondary aging temperature, and secondary aging time were taken as input heat treatment variables and appraised in terms of mechanical property and microstructure. The research study highlights the following outcomes based on the experiments:

1. Taguchi design method is suitable to analyze and optimize the heat treatment parameters as described in the best parameters of hardness and hot tensile test for optimization, with a primary aging temperature of 1120 °C, primary aging time of 3 h, secondary aging temperature of 845 °C, and secondary aging

time of 24 h. The regression model for the hardness of the collected data was approximately 83.72% (R^2_{adj}), which shows that the response can be described by the experimental factors;

2. The predicted hardness values that were obtained using a regression model were consistent with the experimental values as follows: $Y_{HV} = 220 + 0.279(X_1) + 3.880(X_2) - 0.0362 (X_3) + 0.414 (X_4)$;
3. In the Ni-based superalloys grade GTD111, after rejuvenation heat treatment, the gamma prime particle (γ′ or Ni3(Al,Ti)) precipitate was found to be evenly distributed on the gamma (γ). The EDS test showed additional related elements of discontinuous MC carbides (Ta,Ti,W), and the carbide layer was scattered throughout the matrix;
4. The size and shape of precipitates of γ′ and carbide phase was controlled through the heat treatment process. In the optimization condition with the γ′ particle, the characteristics were a coarse square shape and spherical shape of γ′ and fine γ′ precipitate distributed on the parent phase, which affects the mechanical properties of the material.
5. Rejuvenation of GTD111 by heat treatment can restore the mechanical properties and microstructure to their status before their long-term service began when the hot tensile test has a tensile stress of 402 MPa. It was found that high-temperature rupture of the material with suitable conditions caused a micro crack at the position of the carbide, which resulted in the material being brittle.

5. ACKNOWLEDGEMENT

This research supported by the Petch Pra Jom Klao Master's Degree Research Scholarship from King Mongkut's University of Technology Thonburi.

6. REFERENCES

1. A.B. Esanab, V. Ehiaguinab, C. Awosopeb, L. Olatomiwac and D. Egbunea, “Data-based investigation on the performance of an independent gas turbine for electricity generation using real power measurements and other closely related parameters”, ***Data in Brief,*** Vol. 26, (2019), 1-16, https://doi: 10.1016/j.dib.2019.104444.
2. Matthew J. et al. 2002, Superalloy A Technical Guide, ASM International Publications, Ohio, 2000.
3. J.R. Davis, ASM specialty Handbook Nickel Cobalt and Their Alloy, ASM International Publications, Ohio, 2000.
4. M. Perrut, P. Caron, M. Thomas and A. Couret,, “High temperature materials for aerospace applications: Ni-based superalloys and γ-TiAl alloys” , ***Comptes Rendus Physique,*** Vol. 19, (2018), 657-671, https://doi.org/10.1016/j.crhy.2018.10.002.
5. B. Salehnasab, E. Poursaeidi, S.A. Mortazavi and G.H. Farokhian, “Hot corrosion failure in the first stage nozzle of a gas turbine engine”, ***Engineering Failure Analysis,*** Vol. 60, (2016), 316-325, https://doi.org/10.1016/j.engfailanal.2015.11.057.
6. P. Zhangab, J. Lic, X. F. Gongd,Y. Yuana, Y.F. Gua, J.C. Wangb, J.B. Yana and H.F. Yina, “Creep behavior and deformation mechanisms of a novel directionally solidified Ni-base superalloy at 900°C” , ***Materials Characterization,*** Vol. 148, (2019), 201-207, https://doi.org/10.1016/j.matchar.2018.12.023.
7. J.K. Kim, H.J. Park, D.N. Shim and D.J. Kim, “Transient liquid phase bonding of γ′- precipitation strengthened Ni based superalloys for repairing gas turbine components”, ***Journal of Manufacturing Processes,*** Vol. 25, (2017), 60-69, https://doi.org/10.1016/j.jmapro.2016.10.002.
8. A. Rezaie and S.E. Vahdat, “Study of effects of temperature and pressure in HIP process on mechanical properties of nickel-based superalloys”, Materials ***Today Proceeding,*** Vol. 4, (2017), 152-156, https://doi.org/10.1016/j.matpr.2017.01.008.
9. X.W.Jiang, D.Wang, Di.Wang, H.Li and L.H.Lou, “The effect of reheat treatment on microstructure and stress rupture property of a directionally solidified nickel-based superalloy after long-term thermal exposure”, ***Materials Science and Engineering: A,*** Vol. 694, (2017), 48-56, https://doi.org/10.1016/j.msea.2017.03.116.
10. Y.Y. Kaplanskii, E.A. Levashov, A.V. Korotitskiy, P.A. Loginov, Zh. A. Sentyurina and A.B. Mazalov, “Influence of aging and HIP treatment on the structure and properties of NiAl-based turbine blades manufactured by laser powder bed fusion”, ***Additive Manufacturing,*** Vol. 31, (2020), 1-12, https://doi.org/10.1016/j.addma.2019.100999.
11. J. Xua, H. Gruberb, D. Denga, R. Lin, P. Johan and J. Moverarea, “Short-term creep behavior of an additive manufactured non-weldable Nickel-base superalloy evaluated by slow strain rate testing”, ***Acta Materialia,*** Vol. 179, (2019), 142-157, https://doi.org/10.1016/j.actamat.2019.08.034.
12. C.D. Pimenta, M.B. Silva, R.L. de Morais Campos, and W.R. de Campos Junior, “, Application of the taguchi method in the investigation of influential heat treatment factors in steel wires”, ***International Journal of Innovative Research and Development,*** Vol. 8, (2019), 111-114, https://doi:10.24940/ijird/2019/v8/i3/MAR19029
13. H. Terzioglu, “Analysis of effect factors on thermoelectric generator using Taguchi method”, ***Measurement,*** Vol. 149, (2019), 1-10, https://doi: 10.1016/j.measurement.2019.106992.
14. G. kartheek, K. Srinivas and Ch. Devaraj, “Optimization of residual stresses in hard turning of super alloy inconel 718, Materials today: Proceedings”, ***Materials today: Proceedings,*** Vol. 15, (2018), 4592-4600, https://doi.org/10.1016/j.matpr.2017.12.029.
15. A. Kumar, A.R. Ansari, B. N. Roy, and S. Kumar, “Heat treatment parameter optimization using Taguchi technique”, ***International Journal of Scientific Research and Education,*** Vol. 4, (2016), 5965-5974, https://doi.org/10.18535/ijsre/v4i10.07.
16. R.M. Dodo, T. Ause, E.T. Dauda, U. Shehu and A.P.I. Popoola, “Multi-response optimization of transesterification parameters of mahogany seed oil using grey relational analysis in Taguchi method for quenching application”, ***Helicon,*** Vol. 5, (2019), 1-7, https://doi: 10.1016/j.heliyon. 2019.e02167.
17. M. Sobhani, H. Ahmadi, T. Javad and A. Esfahani, “Taguchi optimization of combined radiation/natural convection of participating medium in a cavity with a horizontal fin using LBM”, ***Physica A: Statistical Mechanics and its Applications,*** Vol. 509, (2018), 1062-1079, https://doi.org/10.1016/j.physa.2018.06.056.
18. A.S. Canbolata, A.H. Bademlioglub, N. Arslanoglua and O. Kaynaklia, “Performance optimization of absorption refrigeration systems using Taguchi, ANOVA and grey relational analysis

methods", ***Journal of Cleaner Production,*** Vol. 229, (2019), 874-885, https://doi.org/10.1016/j.jclepro.2019.05.020.

19. J.A. Daleo and J.R. Wilson, "GTD111 alloy material study", ***Journal of Engineering Gas Turbines Power,*** Vol. 120, (2013), 375-382, http:// doi:10.1.1.875.4753&rep=rep1&type=pdf.
20. A. Rezaie and S. E. Vahdata, "Study of effects of temperature and pressure in HIP process on mechanical properties of nickel-based superalloys", ***Materials Today: Proceedings,*** Vol. 4, (2017), 152-156, https://doi.org/10.1016/j.matpr.2017.01.008.
21. Kathleen Mills, ASM Handbook Volume 9 Metallography and Microstructures, ASM International Publications, Ohio, (1985)
22. John N. Dupont,, Welding Metallurgy and weldability of nickel base alloy, A John Wiley and Son Publishing, New Jersey,(2009)
23. F. Long, Y.S. Yoo, C.Y. Jo, S.M. Seo, Y.S. Song, T. Jin and Z.Q. Hu, "Formation of η and σ phase in three polycrystalline superalloys and their impact on tensile properties", ***Materials Science and Engineering A,*** Vol. 527, (2009), 316-369, https://doi.org /10.1016/j.msea.2009.09.016.
24. M. Berahmand and S. A. Sajjadi, "Morphology evolution of γ′ precipitates in GTD-111 Ni-based superalloy with heat treatment parameters, "International Journal of Materials Research", ***International Journal of Materials Research,*** Vol. 104, (2013), 275-280, https://doi.org/10.3139/146.110856.
25. Q. Wang, Z. Li, S. Pang, X Li, C Dong and P. K. Liaw, "Coherent precipitation and strengthening in compositionally complex alloys: A Review", ***Entropy,*** Vol. 20, (2018), 1-23, https://doi.org/10.3390/e20110878.
26. P. Zhanga, Y. Yuan, J. Li, Y.F. Xu, X.L. Song and G.X. Yang, "Tensile deformation mechanisms in a new directionally solidified Ni-ba se superalloy containing coarse γ′ precipitates at 650 °C", ***Materials Science and Engineering A,*** Vol. 702, (2017), 343-349, https://doi.org/10.1016/j.msea.2017.07.025.
27. M. Vinoth Kumar, V. Balasubramanianb and A. Gourav Rao, "Hot tensile properties and strain hardening behavior of super 304HCu stainless steel", ***Journal of Materials Research and Technology,*** Vol. 6, (2016), 116-122, https://doi.org/10.1016/j.jmrt.2016.05.004.
28. A. Dadkhah and A. Kermanpur, 2017, "On the precipitation hardening of the directionally solidified GTD-111 Nibase superalloy: Microstructures and mechanical properties", ***Materials Science and Engineering A,*** Vol. 685, (2017), 79-86, https://doi.org/10.1016/j.msea.2017.01.005.
29. S. A. Sajjadi, S. Nategh, M. Seyed, and M. Zebarjad, "Tensile deformation mechanisms at different temperatures in the Ni-base superalloy GTD-111", ***Journal of Materials Processing Technology,*** Vol. 155-156, (2004), 1900-1904, https://doi.org/10.1016/j.jmatprotec.2004.04.273.
30. L. Zhang, Y. Zhou, X. Jin, X. Du and B. Li, "The microstructure and high-temperature properties of novel nano precipitation-hardened face centered cubic high-entropy superalloys", ***Scripta Materialia,*** Vol. 146, (2018), 226-230, https://doi.org/10.1016/j.scriptamat.2017.12.001.
31. B. G. Choi, I. S. Kim, D. H. Kim and C. Y. Jo, "Temperature dependence of MC decomposition behavior in Ni-base superalloy GTD 111", ***Materials Science and Engineering A,*** Vol. 478, (2008), 343-349, https://doi.org/10.1016/j.msea.2007.06.010.

Persian Abstract

چکیده

این تحقیق یک بهینه سازی و جوان سازی فرآیند عملیات حرارتی برای درجه فوق العاده آلیاژ پایه نیکل GTD۱۱۱ را پس از خدمات طولانی مدت توصیف می کند. متغیرهای عملیات حرارتی پیری بررسی شده در این مطالعه شامل دمای پیری اولیه ، زمان پیری اولیه ، دمای پیری ثانویه و زمان پیری ثانویه است. مواد بدست آمده با استفاده از روش آزمایش تاگوچی آزمایشات تعیین شد تا آزمایش سختی ماده حاصل شود و با آزمایش کشش گرم ، میکروسکوپ الکترونی روبشی و طیف سنجی اشعه X پراکندگی انرژی مشاهده شد. نتایج تجربی نشان داد که در پی بهینه سازی با پارامترهای عملیات حرارتی دمای پیری اولیه ۱۱۲۰ درجه سانتیگراد ، زمان پیری اولیه ۳ ساعت ، دمای پیری ثانویه ۸۴۵ درجه سانتیگراد و زمان پیری ثانویه ۲۴ ساعت چه اتفاقی می افتد. این ماده ، پس از عملیات حرارتی جوان سازی از طریق بهینه سازی با ویژگی های ذرات γ ، دارای یک شکل مربع درشت ، شکل کروی γ' ، و رسوب fine γ خوب توزیع شده بر روی فاز اصلی است ، که بر خصوصیات مکانیکی مواد تأثیر می گذارد. رسوب خوب γ distributed توزیع شده بر روی فاز اصلی ، که بر خصوصیات مکانیکی مواد تأثیر می گذارد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 966-975

International Journal of Engineering

Journal Homepage: www.ije.ir

Modeling and Optimization of Charge Materials Ranges in Converter Furnace with Enhanced Passivation Time in Copper Electrorefining Process: A Mixture Design Approach

M. H. Khazaei Feizabad*[a], G. R. Khayati[a], R. Kafi Hernashki[c], S. M. J. Khorasani[d]

[a] *Department of materials science and Engineering, Shahid Bahonar University of Kerman, Kerman, Iran*
[b] *Research and development center, shahrebabak copper complex, national Iranian copper Industries company, Kerman, Iran*
[c] *Research & Development Center, Shahrebabak Copper Complex, National Iranian Copper Industries Company, Kerman, Iran*
[d] *Senior metallurgical engineer, Process control unit, Khatoonabad copper refinery, Shahrebabak copper complex, National Iranian copper industries company, Kerman, Iran*

P A P E R I N F O

Paper history:
Received 21 September 2020
Received in revised form 26 January 2021
Accepted 13 February 2021

Keywords:
Copper Electrorefining
Passivation Time
Casting Anode
Design of Experiment

A B S T R A C T

In this study, the design of experiments is used to study and model the time of passivation in copper electrorefining as a function of the charge of melting furnace through the preparation of copper casting anodes. As a result of optimization for the proposed optimized anodes, the charge percent values of concentrate (Co), refinery scrap (RS), and non-refinery scrap (NRS) were proposed equals to 69.1, 0.574 and 30.32 (wt.%), respectively. Experimental data confirmed the enhanced passivation time of the proposed anode was 6520 s. Also, it was observed that the molar ratio of As/(Bi+Sb) and Ag/(Se+Te) are the key factors in passivation time. Finally, the relation of passivation time (seconds) with the charge of melting furnace is proposed as : t (s)= - 3728.98 × Co + 4640.00 × RS + 3141.00 × NRS + 17763.27 × Co × RS + 25547.65 × Co × NRS - 1758.00 × RS × NRS. Moreover, adding of As ingot in casting anodes as a dose dependent of non-refinery scrap portion in the input charge of the melting unit can effectively prolong the time of passivation.

doi: *10.5829/ije.2021.34.04a.23*

1. INTRODUCTION

Almost all produced copper by different methods including hydrometallurgy or pyrometallurgy, experience an electrolytic step. The former uses the electrolytic step through the production of pure copper from the impure casting anode by the electrorefining process. While the later uses the electrolytic step through the electrowinning process [1-3].

The electrorefining process involves two general steps. The first step is the anodic electrochemical dissolution of impure copper in an electrolyte containing H_2SO_4 and $CuSO_4$ and the second step is the deposition of pure copper on the cathode. In this process, pure copper is produced and also, valuable impurities are removed from the impure copper as anodic slime [1, 2, 4]. Copper anodes with a purity of 98.5 to 99.5% is used to produce a copper cathode with a purity of more than 99.997% in the electrorefining process [1].

Passivation is one of the most important problems in the industrial electrorefining plant. It refers to a sudden increase in the applied cell voltage. This problem may be due to the formation and accumulation of various nonconductive corrosion products at the anode surface. This passive layer acts as diffusion barriers and prevents the diffusion of copper ions from the anode to the cathode. In this condition, the production capacity reduces and increases energy consumption. It also increases the amount of slime in the electrolyte, which can reduce the quality of the produced cathode. The anode dissolution stops or becomes very slow during the passivation while the electrorefining process is not yet completed. These results cause a lot of energy and work

*Corresponding Author Institutional Email: *m.h.khazaei@eng.uk.ac.ir* (M. H. Khazaei Feizabad)

Please cite this article as: M. H. Khazaei Feizabad, G. R. Khayati, R. Kafi Hernashki, S. M. J. Khorasani, Modeling and Optimization of Charge Materials Ranges in Converter Furnace with Enhanced Passivation Time in Copper Electrorefining Proces: A Mixture Design Approach, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 966-975

loss because the anode has to be removed and melted again [5-13].

In industrial conditions, pure copper anodes cannot be passivated, so anode impurities are the main reason for industrial passivation. These impurities originated from the charges of the melting furnace. When dissolving the anode, some of these impurities appear as slime and the rest dissolve in the electrolyte. Such impurities increases in the electrolyte over time and real-time purification of electrolyte is necessary [1, 5-7].

Due to the electrochemical nature of the electrorefining process, different electrochemical methods are used to study the passivation behavior of the casting anode. Cyclic voltammetry (CV), linear potential sweep (LPS), and galvanostatic or chronopotentiometry (CP) measurement are the most important of these methods. Since the most important cause of passivation is the presence of impurities, the method that is more sensitive to impurities is more suitable for studying of passivation [9, 14]. It is shown that the CP is the most sensitive method and provides the best quantitative data for anode passivation [14]. Also, this technique determines the "passivation time" in an experimental condition that can be used as a criterion for comparison of the anode passivation at the industrial condition.

There are various techniques for optimizing the experimental processes. All of them provide the conditions with the minimum number of tests without a significant impact on the accuracy of the measurement. Also, these techniques provide the possibility of determination of optimum condition. Response surface methodology was proposed employed to design the experiments through the collecting data [15]. It is one of the most promising approaches in the case that the operational factors consider as ingredients of the mixture and the response is a dependent proportions function [16-20]. Low operating costs due to the reduction in the number of required tests and provide the possibility of the simultaneous changes in variables are the main advantages of this approach [21, 22]. To the best of our knowledge, despite extensive studies on the anode passivation, it is suffered from the systematic investigation about the optimization of proposed compounds .

The charge of copper anode casting furnace in the Khatoonabad copper refinery of Iran is supplied from three sources including concentrate (Co), refinery scrap (RS), and non-refinery scrap (NRS), which can be charged in a different ratio. In this study, the effect of each charging source is estimated on the occurrence of passivation using the mixture method as one of the most promising approaches in the design of experiments. Then, the passivation of the casting anode is exanimated by CP.

2. MATERIALS AND METHODS

According to the literatures [6-12], the most important factor affecting the passivation of industrial anodes is their chemical composition. The chemical composition of the anodes is determined at the casting machine, so the input materials of the melting furnace must be studied to evaluate the effect of the chemical composition. In the Khatoonabad copper complex of Iran, the entrance of the smelting furnace includes Co, RS, and NRS. In this case, to change the chemical composition, the input component of the furnace charge must be changed. Therefore, to study the effect of input and determine the appropriate composition to increase the onset time of passivation, the experimental design method can be used.

Since in this study, the total input of the furnace is fixed and only the fraction of each input change, the mixture experiment design method should be used as a powerful method in examining and optimizing the inputs of the smelting unit. This method can be used to determine the optimal charge ratio in the melting furnace to increase the passivation time. According to the limit announced by Khatoonabad copper complex for each of the inputs of the smelting furnace, the limits of the inputs of the furnace are considered according to Table 1.

According to the proposed model by mixed design, the relationship between the passivation time and melting furnace charge can be presented as a polynomial model. It should be noted that the total amount of furnace charge components must be equal to 100. Due to these limitations, the polynomial model can be presented in a standard form. Equation (1) presents the standard model:

$$Y^* = b1X1 + b2X2 + b3X3 + b12X1X2 + b13X1X3 + b23X2X3 + b123X1X2X3 \tag{1}$$

where $Y = Y^* + e$ and X1, X2, and X3 are the selected parameters through modelin.

Design-Expert software is used to design the experiments through the data collection. Table 2 shows 7 experiments that proposed by Design-Expert Software. For each of the experiments, the passivation time of the samples is measured, separately by chronopotentiometry analysis. To estimate the error of analysis, one of the experiments is repeated, randomly (in this study A7).

All samples polish completely to remove the effect of the sample morphology. According to the discussions,

TABLE 1. Range of furnace charge component (Concentrate: Co, Refinery scrap: RS, Non-refinery scrap: NRS)

Charge of furnace	Lower limit	Upper limit
Co	50	85
RS	0	50
NRS	0	50

TABLE 2. Experimental design performed in the current study

Run	Co	RS	NRS
A1	50	25	25
A2	50	0	50
A3	85	15	0
A4	50	50	0
A5	85	0	15
A6	67.5	32.5	0
A7	85	0	15

the chemical composition is the most effective factor in the passivation process, so to determine the exact chemical composition of the prepared samples in the previous step, the quantometer analysis is used .

The galvanostatic test uses to investigate the electrochemical behavior of prepared samples. In this technique, the potential fluctuations of the sample at constant current density are measured. A platinum electrode is used as a cathode to reduce the cathode effect through the process. Due to its high resistance to the chemical reactions, platinum does not affect the process and uses only as a current transfer in the electrochemical circuit. All prepared samples in the previous steps are used as anodes. The anode surface in contact with the electrolyte solution is 1 cm^2. Other surfaces of all samples are coated to ensure that other surfaces of the sample are not affected by the test.

The electrolyte uses in this research is the industrial electrolyte of the Khatoonabad copper complex, which has a combination of 180 g/l sulfuric acid and 45 g/l copper. Additives of the electrolyte are glue, thiourea, and chloride ions. The amounts of glue and thiourea are 95 g/ton and 110 g/ton of cathodic copper, respectively. Also, the amount of chloride ion is kept constant to be about 40 ppm in the electrolyte. To simulate the electrolyte motion at the industrial electrorefining cell, a magnetic stirrer device is used at the speed of 100 rpm. Also, similar to the industrial conditions, the electrolyte temperature is kept constant at 62 ºC during the process. Moreover, the samples are tested for 7200 s at a current density of 220 mA/cm^2. In this research, an EG & G 263A potentiostat/galvanostat device is used. To accurately investigate the potential fluctuations, a ternary arrangement of electrodes has been used. This arrangement consists of three electrodes. The auxiliary electrode is the platinum cathode, the working electrode is the copper sample prepared as the anode, and the third electrode is the reference electrode. To increase the accuracy of potential measurement, the sample potential should be measured relative to a reference electrode with a constant potential i.e., a Hg/Hg_2Cl_2 calomel in this study .

Given the engineering knowledge and experience, the Khatoonabad copper refinery uses a mental model based on the operator's skills to add the As ingot through the melting and preparation of casting anodes. In this pattern, arsenic added proportional to the portion of NRS in the charge of melting furnace, this mental pattern significantly affects the processing efficiency and is dependent variables. Accordingly, it is not considered as independent parameters.

3. RESULTS AND DISCUSSION

3. 1. Galvanostatic Study The prepared samples, which have an industrial composition, are subjected to a current density of 220 mA/cm^2. The reason for the use of such a high density that accelerates the dissolution rate is the limitations of operating time through the electrochemical tests. Figure 1 shows the chronopotentiometry diagrams of casted anodes samples. All diagrams have three distinct regions including the active zone, the pre-passive zone, and the passive zone.

The passive mechanism can be summarized as that when the casting anode is dissolved, the anode slime begins to release at the anode surface, the active regions decrease, and the current density increases. Also, the concentration of copper at the anode surface increases. When the concentration of copper exceeds the saturation at the electrode surface, it begins to precipitate on the surface of the anode and inside the slime. As a result, slime and copper sulfate cover the surface of the anode. The formation of a sulfate layer on the surface of the electrode allows the oxide layer, which is only stable under certain conditions, to form below the surface of the sulfate layer. But with the formation of the sulfate layer, the copper concentration of the anode surface decreases, which can lead to the re-dissolution of the sulfate layer. If there are no suitable conditions for the stability of the oxide layer when the sulfate layer is dissolved, this layer will also dissolve in the electrolyte. However, if the potential of the oxide layer is sufficiently increased after the sulfate layer is dissolved, the oxide layer does not dissolve and the sample becomes passive. Due to this passive mechanism, any factor that disrupts these phenomena increases the passivation time. In some of the samples, the active dissolution and passivation process repeated several times that is the result of the separation of the anode slime.

In the active region, the dissolution of the metal anode occurs at the high current densities. Copper dissolution with the following reaction can create a passive stable layer of Cu_2O due to the low potential of reaction (2):

$$Cu+H_2O \leftrightarrow Cu_2O+2H^++2e^- \quad E_{CuO/Cu}=-34 \text{ mV (vs MSE} \quad (2)$$

Although there is no change in the appearance of potential in the active region and it is constant, in this

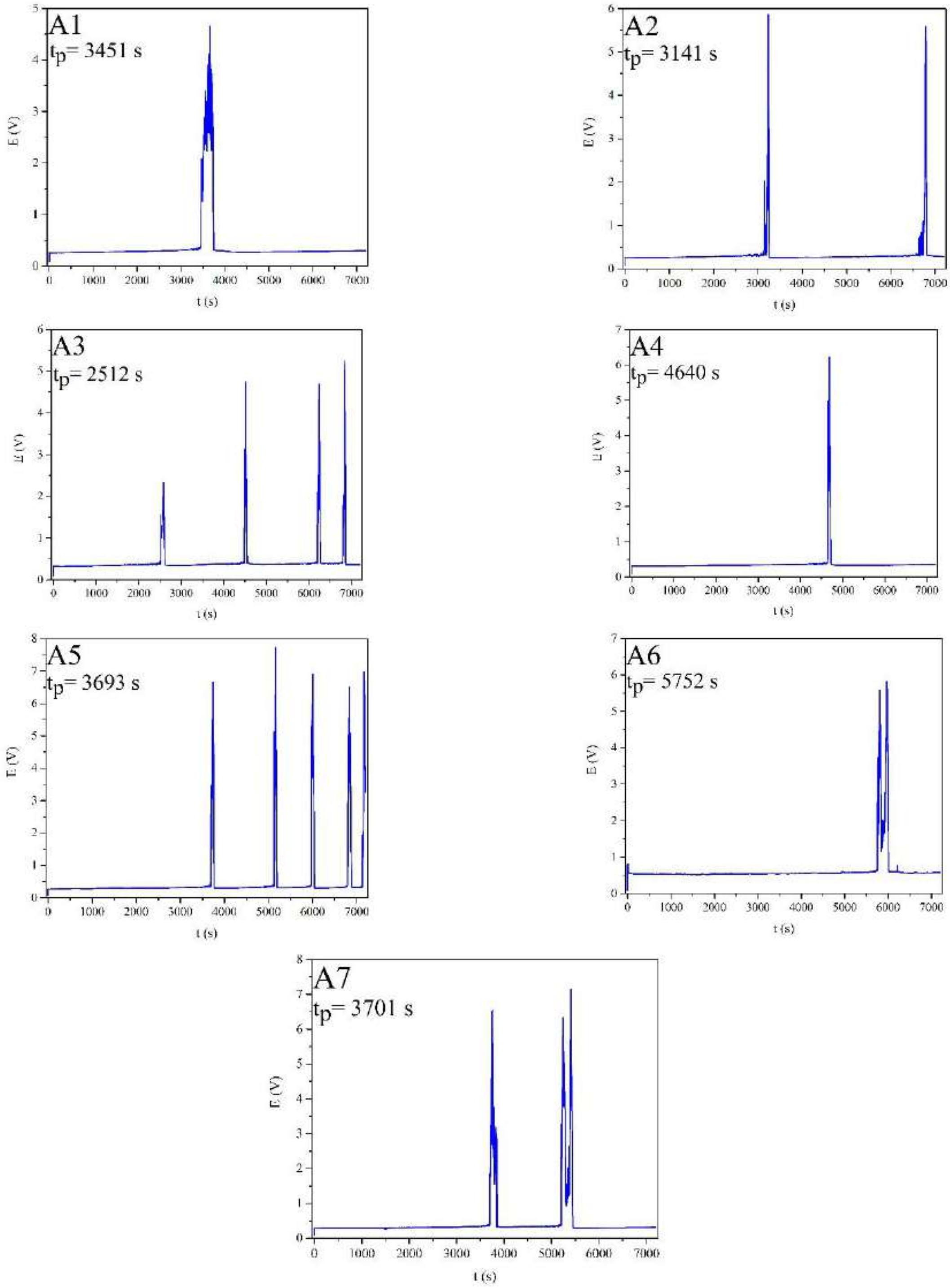


Figure 1. The chronopotentiometry analysis of prepared samples

region the potential increases with a small slope relative to the time. It should be noted that applying a constant current density to dissolve copper does not mean a fixed corrosion rate during the test. Corrosion rates can change during the following condition [6]:

1. The presence of different reactions that consume the currents:

$$Cu \rightarrow Cu^{2+} + 2e^- \quad (3)$$

$$Cu \rightarrow Cu^{+} + e^- \quad (4)$$

The formation of Cu^{+} ions dissolves twice the amount of copper relative to Cu^{2+} at a constant current. The equilibrium between the concentrations of these two copper ions in the solution is expected to be stable at 25ºC based on the following relationship (Equation (5)):

$$\frac{[Cu^{2+}]}{[Cu^{+}]^{2}} = 10^{6.9} \tag{5}$$

Cu^{+} consumption increases with Cl^{-} or anode impurities such as Se, and increases the rate of Cu^{+} formation, which can accelerate the copper dissolution;

2. The surface of the anode is covered with various corrosion products, which reduces the active level of the anode and leads to an increase in the current density at the active levels;
3. The presence of different phases in the anode microstructure, which is quite conceivable due to the different chemical composition of the industrial anodes, can lead to a different corrosion rate in the samples. The presence of impurities causes more current consumption;
4. Increasing the surface roughness of the anode can lead to an increase in the anode surface during the corrosion, thus reducing the density of the applied current. However, in this study, an attempt has been made to reduce the effect of this factor as much as possible by polishing the surface of the samples.

In the pre-passive zone, anode dissolution and deposition of corrosion products occur sequentially, resulting in the formation of a passive layer with temporary stability on the anode surface. Given the gradual increase in potential, it is expected that the dominant response at this stage will be the formation of CuO as a product with less stability and greater dissolution. The reaction is as follows (Equation (6)):

$$Cu + H_2O \leftrightarrow CuO + 2H^{+} + 2e^{-} \quad E_{CuO/Cu} = -34\text{ mV (vs MSE)} \tag{6}$$

This oxide is gradually formed in the cavities of the passive layer Cu_2O. Cu^{2+} saturation at the surface of the passive metal causes the formation of $CuSO_4$. Increasing the level of Cu^{+} ions in the surface causes the following reaction (Equation (7)) and the formation of copper powder.

$$2Cu^{+} \leftrightarrow Cu^{2+} + Cu \tag{7}$$

At this condition, copper sulfate and copper powder cover the surface of the copper oxide layer. Thus, anode-level products generally include copper powder, copper sulfate, and copper oxide (CuO and Cu_2O), as well as the other additional components due to the impurities of the anode compound.

Figure 2 shows a free corrosion diagram of a copper-water system at 25ºC. Various parameters such as temperature, corrosion formation on the surface and their accumulation near the surface, lack of penetration due to low excitation, and the effect of additives on the surface can change the pH of the reaction near the surface. In addition, the high current density of 220 mV/cm^2 determines the dominant reaction of the trans-passive region of the Purbe diagram, while this graph generally explains the active-passive region. All of these parameters cause a sudden change in pH near the surface where different reactants such as Cu_2O, Cu^{+}, and impurities such as Sb_2O_3 and Bi_2O_3 consume the acid at the joint level of the metal with the solution. Finally, according to Purbe's diagram (Figure 2), increasing the pH leads to the stability of Cu_2O. In the passive region, the formation of a stable Cu_2O layer is the predominant reaction. It should be noted that in this area, the reaction is accelerated (Equation (8)) in the presence of oxygen.

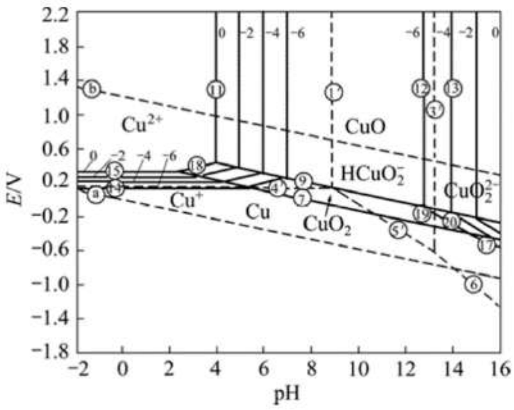


Figure 2. Purbe's diagram of Cu-H_2O system at 25˚C [6]

$$1/2\ O_2 + Cu_2O \rightarrow 2CuO \tag{8}$$

After the passive region, several active-passive unstable steps are observed again. This can be due to the inadequacy of passive products on the common metal-soluble surface. Also, the lack of sufficient adhesion of the products can lead to the separation of the passive layer and the creation of the next active-passive steps.

Figure 3 shows the first and second regions of the chronopotentiometry diagram, the active and pre-passive regions of the tested samples. Although the potential changes in the chronopotentiometry chart appear to be constant in the first and second regions, in fact, the potential in these areas is increasing with a slight slope. This increase in potential could be due to the formation of corrosion products and the release of sludge at the anode level, which leads to an increase in potential. Observations show that increasing the slope of the first zone eventually leads to accelerated passivation. Sample 6 has the least slope changes between the samples and the longest passivation time. The reason for this can be stated in the arsenic value of this sample. Increasing arsenic reduces the adhesion of corrosion products and separates them from the surface, which can reduce the slope of the potential increase and increase the passivation time.

3. 2. Modeling of Passivation Time The time of passivation, i.e., the experiment design response, were added to Figure 2. Equation (9) indicates the proposed model for determining the time of passivation in terms of melting load furnace.

$$t = -3728.98 \times Co + 4640.00 \times RS + 3141.00 \times NRS + 17763.27 \times Co \times RS + 25547.65 \times Co \times NRS - 1758.00 \times RS \times NRS \quad (9)$$

ANOVA analysis is one of the most widely used methods of analysis in statistical studies. This analysis is a powerful tool for estimating the statistical models due to the distribution of experimental data obtained in this method. In this study, the model provided for passivation time has a P-value equal to 0.0037, which indicates that the model is acceptable. Table 3 shows ANOVA analysis for the passivation time of prepared samples.

According to the proposed software model, the predicted values of the passivation time of samples with different furnace load combinations are shown in Figure 4.

The chemical compositions of prepared samples are abbreviated in Table 4. As it turns out, the chemical composition of the samples is very wide, and it is clear that it has a complex relationship with the furnace loads. Because of the complexity of the relationship between furnace input and anode chemical composition, the experimental design method can be a good way to examine and process the data.

TABLE 3. the ANOVA analysis for passivation time of prepared samples

Source of variation	Sum of Squares	Df	Mean Square	F-Value	P-Value	Significant
			Response Y_1			
Model	6.740E+006	5	1.348E+006	42125.84	0.0037	Significant
Linear Mixture	2.391E+006	2	1.196E+006	37361.58	0.0037	
AB	3.157E+006	1	3.157E+006	98645.33	0.0020	
AC	3.949E+006	1	3.949E+006	1.234E+005	0.0018	
BC	1.288E+005	1	1.288E+005	4024.17	0.0100	
Pure Error	32.00	1	32.00	12.06	0.1785	
Cor Total	6.740E+006	6				

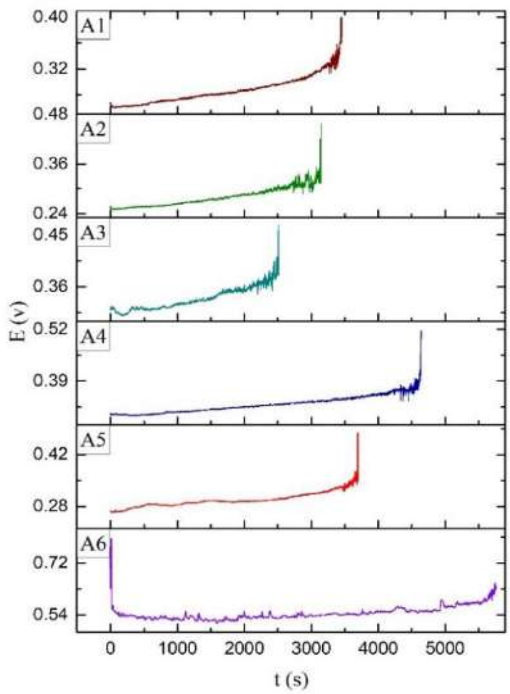


Figure 3. Illustration of the active and pre-passive region in various samples

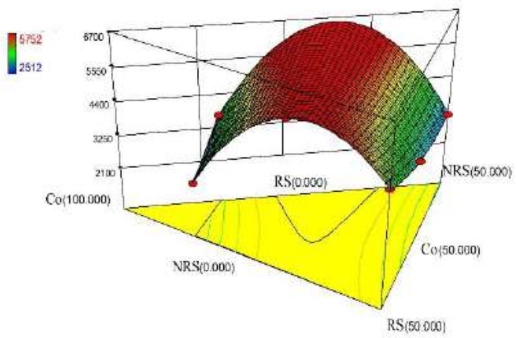


Figure 4. 3D profile of passivation time versus the load charge of melting furnace

The highest passivation time in the experiments performed according to Figure 3 is related to the A6 sample. As shown in Table 4, this sample has the highest amount of As. Research has shown that the presence of As in the compound can reduce the adhesion of the slime layer and cause the layer to separate from the anode surface [11]. Also, As is an acid-forming element based (Equation (10)). Decreasing the pH of the electrolyte leads to the instability of the copper oxide and its

TABLE 4. Chemical combination of prepared samples

	%Cu	P	S	Te	As	Se	Sb	Cd	Cr	Pb	Si	Bi	Ag	Zn	Ni	Co	Fe	$\%O_2$	Au	Sn	Mn
A1	99.64	0.5>	12.53	50.0	436	124	211	0.5>	0.5>	202	7.69	8.0	239	36.77	260.5	2.43	83.86	0.1	7.44	1.43	1.23
A2	99.55	<0.5	8	44	464	119	184	<0.5	<0.5	200	<0.5	8.98	244	37	278	1.57	69	0.19	7.94	<0.5	<0.5
A3	99.63	<0.5	7.39	39	451	122	312	<0.5	<0.5	165	<0.5	11	209	23	263	0.5	49	0.18	7.10	3.77	<0.5
A4	99.66	<0.5	19.5	48.4	451	121	179	<0.5	<0.5	209	<0.5	7.22	230.7	47.5	269	2.87	35.01	0.17	7.31	0.87	<0.5
A5	99.66	0.5>	3.95	32.0	419	98	190	0.5>	0.5>	136	0.5>	7.36	226.3	23.47	223.9	0.5>	62.09	0.12	7.60	8.33	0.5>
A6	99.66	<0.5	2.76	24	497	68.1	336	<0.5	<0.5	314	<0.5	11.3	197.4	20.90	458.6	<0.5	23.3	0.13	6.50	43.16	<0.5

dissolution in the electrolyte, which increases the passivation time.

$$As_2O_3 + 3H_2O \rightarrow H_2AsO_4 + 2H^+ + 2e \quad (10)$$

But as the results of this study showed, As is not the only determining factor in passivation. For example, two samples, A3 and A4, have the same As value with a large difference in passivation time. In other words, the shortest and longest passivation times were respectively belonged to A3 and A4 samples. These two samples even have almost the same amount of Ni, but the important difference between these two samples is the amount of Sb and Fe that higher in A3. Sample A3, after the A6 sample with the highest value of As, has the highest Sb content. Sb and Bi are the elements that can be present in the electrolyte in the form of suspended slime. Also, these elements are (Equations (11) and (12)) the acid-consuming elements that can ultimately lead to the stability of the copper oxide layer and reduce the passivation time [23].

$$Bi_2O_3 + 6H^+ \rightarrow 2Bi^{3+} + 3H_2O \quad (11)$$

$$2Sb^{3+}+O_2+4H^+=2Sb^{5+}+2H_2O \quad (12)$$

Based on the results of this research, it is possible to find a suitable relationship between the process of passivation and certain ratios of alloying elements. The most important ratio of impurities that can affect passivation is the molar arsenic ratio to the sum of Bi and Sb. This ratio can be used to compare the samples to each other. The lower passivation tendency is observed in anods with the higher values of this molar ratio.

Table 5 shows the calculated ratios of the tested samples. The highest molecular ratio of As/(Bi+Sb) belongs to the A4 sample, which has a relatively long passivation time of 4640 s. A2, A5, and A3 have the following values of this ratio, the values of which are also close to each other. Also, A3, in which the value of the ratio is 2.3, has the shortest passivation time. It is observed that, except for A6, the passive time is strongly dependent on this ratio. A6 has the lowest value of this ratio however shows the longest passivation time. The reason for this could be related to the Ag/(Te+Se) ratio in

TABLE 5. The molar ratio of Ag/(Se+Sb) and As/(Bi+Te) of prepared samples

Sample	As/(Bi+Sb)	Ag/(Se+Te)
1	3.285	1.129
2	3.984	1.221
3	2.301	1.046
4	4.000	1.118
5	3.504	1.406
6	1.741	2.357

addition to the effect of As on the separation of the product layer from the anode level. As shown in Table 5, the value of this ratio in A6 is much higher than the others.

It is worth noting that silver is an effective factor for the rapid removal of Se and Te, which, by forming $Ag_2(Se+Te)$, transfers these elements to the anodic slimes and prevents them from increasing their contents in the electrolyte. Decreasing this ratio, in the case of reducing the silver content in the input, activates the secondary reaction of removing Se and Te with significantly slower kinetics in which copper is replaced by silver, and produces $Cu_2(Se+Te)$. In this condition, due to the slower kinetics than the previous reaction, causes an increase in Se content in the electrolyte. Consequently, the molar ratio of Ag/(Te+Se) is an effective factor for passivation. However, it should be noted that the role of Ag/(Se+Te) molar ratio on passivation is far less than As/(Bi+Sb).

Oxygen content is the next effective factor in the passivation of the anode. Since all samples have the same oxygen content, it is incorrect to address this factor separately. Lead and nickel are the two factors that contribute to the intensification of passivity. Both of them are activated in the presence of oxygen, and their content in A6 is significantly higher than the others. However, the longer the passive time of this sample indicates the greater the effect of the arsenic element, which is higher than in others. Iron is the next element influencing passivation behavior. Since A1 with the highest iron content experienced a relatively long period of passivation, this impurity appears to have had little effect in the study area. This indicates that in addition to the high arsenic value of A6, the higher Ag/(Se+Te) ratio along with the lower iron content, could be one of the reasons for a long time of passivation. Because A6 had far less iron content than A1 with the least time spent.

Overall, the results of this study have shown that in industrial anodes, due to the simultaneous presence of high impurities and the interaction of these elements, the results of the times of passivation have a complex relationship with the chemical composition and cannot be easily relied on. The element established a relationship with the time of passivation. This shows the importance of using the design of the experiment in this research.

3. 3. Validation of Proposed Optimum Condition

Since the passivation time in copper anodes must be increased to increase the efficiency of the electrorefining

TABLE 6. Various combination of charge load in the melting unit by as maximum as possible time of passivation

Number	Co	RS	NRS	t (sec)	Desirability	
1	63.125	0.000	36.875	6283.48	1.000	
2	74.500	0.000	25.500	6159.04	1.000	
3	67.500	32.500	0.000	5752	1.000	
4	66.188	29.375	4.437	5818.52	1.000	
5	64.239	17.386	18.375	5913.83	1.000	
6	64.798	35.202	0.000	5864.42	1.000	
7	65.633	22.491	11.877	5874.96	1.000	
8	64.281	25.048	10.671	5840.73	1.000	
9	66.534	26.462	7.004	5824.38	1.000	
10	73.306	0.000	26.694	6296.31	1.000	
11	63.763	26.115	10.121	5824.41	1.000	
12	63.943	36.057	0.000	5878.37	1.000	
13	70.412	8.294	21.294	6105.48	1.000	
14	68.427	5.111	26.462	6319.38	1.000	
15	67.365	28.137	4.499	5779.45	1.000	
16	67.856	14.162	17.982	6011.06	1.000	
17	69.770	14.614	15.616	5910	1.000	
18	67.194	25.451	7.355	5811.58	1.000	
19	71.075	11.336	17.589	5930.51	1.000	
20	60.234	39.675	0.091	5816.27	1.000	
21	69.511	3.526	26.963	6363.85	1.000	
22	62.726	23.015	14.259	5787.06	1.000	
23	65.128	28.157	6.715	5838.27	1.000	
24	62.317	23.736	13.947	5760.23	1.000	
25	63.451	8.817	27.732	6039.81	1.000	
26	62.899	16.302	20.799	5854.67	1.000	
27	62.362	3.194	34.444	6092.7	1.000	
28	69.100	0.574	30.326	6518.68	1.000	Selected
29	69.717	1.962	28.322	6432.89	1.000	
30	66.294	30.784	2.922	5812.38	1.000	
31	67.625	19.073	13.301	5897.07	1.000	
32	70.944	8.583	20.473	6056.12	1.000	

process, the design-expert software examined the conditions for increasing the passivation time by changing the furnace load in the melting plant. Based on the proposed model, the software has proposed different optimal combinations for the furnace load to increase the passivation time, assuming that the chemical composition of the furnace load components (concentrate, refinery scrap, and non-refinery scrap) is constant. As shown, all compounds have similar passivation time close to each other. According to the results provided in Table 6, combination of 28 of this table has the longest passive time, which is equal to 6518 s. Also, the proposed composition of this sample is equal to 69.1 wt.% of concentrate, 0.574 wt.% of refinery scrap, and 30.326 wt.% of non-refinery scrap.

Figure 5 shows the chronopotentiometry diagrams of the proposed anode composition. It can be concluded that this sample has high passivation time in contrast with the other samples. The passivation time of this sample is 6520 s, which is close to the software estimated passivation time of 6518 s.

The chemical composition of the optimum anode is shown in Table 7. It seems that the higher As dopant in the melting unit proportion to the amount of NRS can effectively suppress the higher potential of passivation in the electrorefining cell.

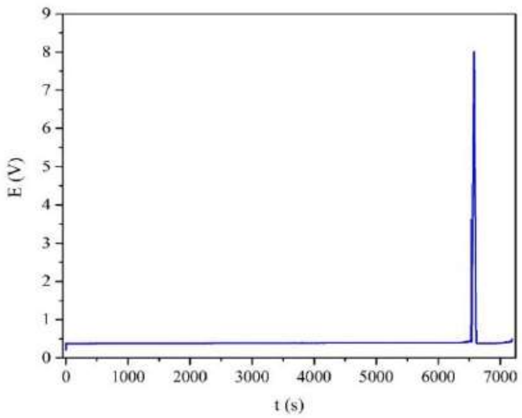


Figure 5. Chronopotentiometry diagrams of the proposed anode at the optimal composition

TABLE 7. The chemical composition of the optimum anode

	%Cu	P	S	Te	As	Se	Sb	Cd	Cr	Pb	Si	Bi	Ag	Zn	Ni	Co	Fe	%O_2	Au	Sn	Mn
Optimal condition	99.66	<0.5	2.76	24	515	70.1	342	<0.5	0.5>	310	<0.5	113	220	21.36	501	<0.5	30.2	0.14	7.20	44.17	<0.5

4. CONCLUSION

The present study shows that the chemical composition of the anode has a significant effect on the process of anode passivation. The chemical composition of the anode is determined in the casting stage, i.e., a function of the charging of the melting furnace. Accordingly, if the chemical composition of the furnace load components is constant, which includes concentrate, refinery scrap, and non-refinery scrap, the optimal fraction of the furnace load components can be obtained by the design of experiment to achieve the maximum time of anode passivation. According to the operational scope of the study, the optimal combined charge composition is 69.1% concentrate, 0.574% refinery scrap, and 30.326% non-refinery scrap to achieve of 6520 seconds passivation time by 220 mA/cm^2 current density. The results of this study also showed that the As/(Bi + Sb) ratio is a good criterion for comparing the passivity of the samples, and higher values of this ratio of 4 mean an increase in the start time of the passivation. Also, the higher value of this ratio can effectively compensate the deteriorating effect of non-refinery scrap impurities in the charge of the melting furnace.

5. REFERENCES

1. Schlesinger, M.E., King, M.J. and Davenport, W.G., "Extractive metallurgy of copper, Elsevier, (2011), doi: 10.1016/C2010-0-64841-3.
2. Jackson, E., "Hydrometallurgical extraction and reclamation, Ellis Horwood Chichester, Vol. 204, (1986), doi: 10.1002/aic.690330923.
3. Li, L., Li, H.-j., Qiu, S.-w. and Wang, H., "Effect of additives on anode passivation in direct electrolysis process of copper—nickel based alloy scraps", ***Journal of Central South University***, Vol. 25, No. 4, (2018), 754-763, doi: 10.1007/s11771-018-3780-1.
4. Jarjoura, G. and Kipouros, G.J., "Electrochemical studies on the effect of nickel on copper anode passivation in a copper sulphate solution", ***Canadian Metallurgical Quarterly***, Vol. 45, No. 3, (2006), 283-294, doi: 10.1179/cmq.2006.45.3.283.
5. Safizadeh, F. and Ghali, E., "Monitoring passivation of cu–sb and cu–pb anodes during electrorefining employing electrochemical noise analyses", ***Electrochimica Acta***, Vol. 56, No. 1, (2010), 93-101, doi: 10.1016/j.electacta.2010.09.046.
6. Safizadeh, F. and Ghali, E., "Electrochemical noise of copper anode behaviour in industrial electrolyte using wavelet analysis", ***Transactions of Nonferrous Metals Society of China***, Vol. 23, No. 6, (2013), 1854-1862, doi: 10.1016/S1003-6326(13)62670-9.
7. Moats, M.S. and Hiskey, J.B., "The role of electrolyte additives on passivation behaviour during copper electrorefining", ***Canadian Metallurgical Quarterly***, Vol. 39, No. 3, (2000), 297-306, doi: 10.1179/cmq.2000.39.3.297.

8. Moats, M.S., Hiskey, J.B. and Collins, D.W., "The effect of copper, acid, and temperature on the diffusion coefficient of cupric ions in simulated electrorefining electrolytes", ***Hydrometallurgy***, Vol. 56, No. 3, (2000), 255-268, doi: 10.1016/S0304-386X(00)00070-0.

9. Minotas, J.C., Djellab, H. and Ghali, E., "Anodic behaviour of copper electrodes containing arsenic or antimony as impurities", ***Journal of Applied Electrochemistry***, Vol. 19, No. 5, (1989), 777-783, doi: 10.1007/BF01320654.

10. Bounoughaz, M., Manzini, M. and Ghali, E., "Behaviour of copper anodes containing oxygen, silver and selenium impurities during electro-refining", ***Canadian Metallurgical Quarterly***, Vol. 34, No. 1, (1995), 21-26, doi: 10.1179/cmq.1995.34.1.21.

11. Abe, S., Burrows, B. and Ettel, V., "Anode passivation in copper refining", ***Canadian Metallurgical Quarterly***, Vol. 19, No. 3, (1980), 289-296, doi: 10.1179/cmq.1980.19.3.289.

12. Chen, T.T. and Dutrizac, J.E., "A mineralogical study of the deportment and reaction of silver during copper electrorefining", ***Metallurgical and Materials Transactions B***, Vol. 20, No. 3, (1989), 345-361, doi: 10.1007/BF02696987.

13. Sędzimir, J. and Gumowska, W., "Influence of electrolysis variables on the passivation time of copper anodes in copper electrorefining", ***Hydrometallurgy***, Vol. 24, No. 2, (1990), 203-217, doi: 10.1016/0304-386X(90)90087-I.

14. Cheng, X. and Hiskey, J.B., "Fundamental studies of copper anode passivation during electrorefining: Part i. Development of techniques", ***Metallurgical and Materials Transactions B***, Vol. 27, No. 3, (1996), 393-398, doi: 10.1007/BF02914903.

15. Box, G.E.P. and Wilson, K.B., "On the experimental attainment of optimum conditions", ***Journal of the Royal Statistical Society: Series B (Methodological)***, Vol. 13, No. 1, (1951), 1-38, doi: 10.1111/j.2517-6161.1951.tb00067.x.

16. Khaskhoussi, A., Calabrese, L., Bouhamed, H., Kamoun, A., Proverbio, E. and Bouaziz, J., "Mixture design approach to optimize the performance of tio2 modified zirconia/alumina sintered ceramics", ***Materials & Design***, Vol. 137, (2018), 1-8, doi: 10.1016/j.matdes.2017.10.010.

17. Khazaei Feizabad, M.H., Sarvestani, E. and Khayati, G.R., "Modeling and optimization of chemical composition of nano/amorphous fea.Nib.Nbc.Zrd alloy prepared via high-energy ball milling with enhanced soft magnetic properties; a mixture design approach", ***Journal of Alloys and Compounds***, Vol. 841, No., (2020), 155646, doi: 10.1016/j.jallcom.2020.155646.

18. Zhang, D., Zhang, Z., Cheng, T. and Zhao, X., "Multi-factorial analysis on vault stability of an unsymmetrically loaded tunnel using response surface method", ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 11, (2019), 1570-1576, doi: 10.5829/IJE.2019.32.11B.08.

19. Mahmood Ali, S., "Optimization of centrifugal casting parameters of alsi alloy by using the response surface methodology", ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 11, (2019), 1516-1526, doi: 10.5829/IJE.2019.32.11B.02.

20. Yousefi, M., Safikhani, H., Jabbari, E., Yousefi, M. and Tahmsbi, V., "Numerical modeling and optimization of respirational emergency drug delivery device using computational fluid dynamics and response surface method", ***International Journal of Engineering***, ***Transactions B: Applications***, Vol. 34, No. 2, (2021), 547-555, doi: 10.5829/IJE.2021.34.02B.28.

21. Eriksson, L., Johansson, E. and Wikström, C., "Mixture design—design generation, pls analysis, and model usage", ***Chemometrics and Intelligent Laboratory Systems***, Vol. 43, No. 1-2, (1998), 1-24, doi: 10.1016/S0169-7439(98)00126-9.

22. Muteki, K., MacGregor, J.F. and Ueda, T., "Mixture designs and models for the simultaneous selection of ingredients and their ratios", ***Chemometrics and Intelligent Laboratory Systems***, Vol. 86, No. 1, (2007), 17-25, doi: 10.1016/j.chemolab.2006.08.003.

23. Petkova, E.N., "Mechanisms of floating slime formation and its removal with the help of sulphur dioxide during the electrorefining of anode copper", ***Hydrometallurgy***, Vol. 46, No. 3, (1997), 277-286, doi: 10.1016/S0304-386X(97)00024-8.

Persian Abstract

چکیده

در این تحقیق برای مطالعه و مدلسازی زمان مورد نیاز برای پسیو شدن آند در فرایند الکتروریفاینینگ مس، برحسب بار کوره ذوب ریخته‌گری آند، از روش طراحی آزمایش استفاده شده است. در نتیجه بهینه‌سازی، نسبت کنسانتره (Co)، قراضه پالایشگاهی (RS) و قراضه غیرپالایشگاهی (NRS) بار کوره، برای آند بهینه به ترتیب ۶۹/۱، ۰/۵۷۴ و ۳۰/۳۲ درصد وزنی تعیین گردید. نتایج آزمایشگاهی، افزایش زمان مورد نیاز برای پسیو شدن را تا ۶۵۲۰ ثانیه برای آند بهینه تایید کرد. همچنین مشاهده گردید که نسبت‌های مولی As/(Bi+Sb) و Ag/(Se+Te) عاملی کلیدی در زمان پسیو شدن آند هستند. در نهایت رابطه زمان پسیو شدن (برحسب ثانیه) براساس کسر اجزا بار کوره ذوب به صورت زیر پیشنهاد شد: $t\ (sec) = -3728.98 \times Co + 4640.00 \times RS + 3141.00 \times NRS + 17763.27 \times Co \times RS + 25547.65 \times Co \times NRS - 1758.00 \times RS \times NRS$. همچنین مشاهده گردید که اضافه کردن شمش As در بار ورودی واحد ذوب ریخته‌گری آند، که مقدار آن وابسته به کسر قراضه غیرپالایشگاهی است، می‌تواند به طور موثری زمان مورد نیاز برای پسیو شدن را افزایش دهد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 976-986

International Journal of Engineering

Journal Homepage: www.ije.ir

A Preliminary Field Study of Antifouling Paint Perfomance After Short Exposure in Mandara Bali, Indonesia

G. Priyotomo*[a], L. Nuraini[a], H. Gunawan[b], J. Triwardono[a], S. Sundjono[a], S. Prifiharni[a]

[a] Research Center for Metallurgy and Material, Indonesian Institute of Sciences, Banten, Indonesia
[b] Research Center and Development of Roads and Bridge, Indonesia's Ministry of Public Works and Housing, Bandung, Indonesia

PAPER INFO

Paper history:
Received 09 June 2020
Received in revised form 13 November 2020
Accepted 08 December 2020

Keywords:
Antifouling Paint
Anticorrosion Paint
Biofouling
Weight Loss
Corrosion
Salinity

ABSTRACT

Antifouling paints are applied to prevent the growth of marine biofouling. In Indonesia, that paint is widely used for ship which commonly used copper-based biocide. In fact, there is no or little comprehesive studies on antifouling paint in Indonesia compared to other tropical countries. In this study, the evaluation of the performance for antifouling paint was carried out where anticorrosion paint and bare steel were also studied as references. The measurement of corrosion rate on steel was conducted by weight loss method. The panels containing specimens were exposure up to 1-month for immersion in different depth of sea up to 3 meters. Seawater parameters consisting of temperature, pH, salinity, conductivity and dissolved oxygen were measured as well as coating properties. The results showed both surfaces of anticorrosion paint and steel specimens covered by biofouling, but not on antifouling paint. There also is not much different in antifouling paint properties before and after exposure in various depth of sea. The reduction of thickness for antifouling paint is apparently predominant to be affected by sea current. The magnitude of corrosion rate on bare steel is almost the same in different depth of sea which took place due to the effect of dissolved oxygen and biofouling. In the future, the comparison of the paints perfomance all local regions is necessary to be conducted in all local regions of the Indonesia.

***doi**: 10.5829/ije.2021.34.04a.24*

NOMENCLATURE

CR Corrosion rate (mpy)	*Corrosion rate, Mils per year (mpy)*	A	area of sample (cm^2)
W	Weight loss (mg)	T	Time of exposure (hour)
D	metal density (g/cm^3)		

1. INTRODUCTION

Biofouling is the term to define undesired plants and animals which attach and grow on submerged artificial static or mobile in seawater. The attachment and colonization of life organisms has a deleterious impact on most offshore submerged structures [1-3]. Moreover, the biological process due to biofouling metabolism could yield detrimental effects on ships such as decreasing the time frequency of dry-docking operation, increasing the frictional resistance, and losing of corrosion resistance [4-7]. The reduction of corrosion resistance could decrease mechanical strength and life time of structure particularly in marine environment [8, 9]. The safety problem is also considered when the decrease of structure stability and concealment structural defects takes place due to the aggressiveness of biofouling growth [9, 10]. Mostly surface temperature, water flow and salinity are predominantly to take the essential part of an activity of biofouling [11] where the other parameters are also considered such as pH and oxygen concentration to stimulate the growth of biofouling in marine water [12]. In tropical region, the rapid growth of marine biofouling occurs mostly due to high seawater temperature and high salinity [13, 14] compared in subtropical region. The growth of attached marine biofouling tends to increase rapidly in low ocean current rates and vice versa [12]. Indonesia has tropical climates, where the seasonal

*Corresponding Author Email: gadangp@gmail.com (G. Priyotomo)

Please cite this article as: G. Priyotomo, L. Nuraini, H. Gunawan, J. Triwardono[a], S. Sundjono, S. Prifiharni, A Preliminary Field Study of Antifouling Paint Perfomance After Short Exposure in Mandara Bali, Indonesia, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 976-986

change is relative stable annually for surface temperature and high salinity in marine environment. Therefore, the growth of tropical marine biofouling continues without disruption in this country.

Commonly the primary mitigation for the severity of biofouling is to utilize antifouling (AF) paint on marine structure application. The mechanism of biocide compound such copper (Cu) releasing from AF paint could minimize the attachment of biofouling on the structure [2]. In addition, the toxic function of copper compound as main additive of AF paint is to prevent the growth of certain biofoulings such as mollucs, algae, bacteria and so on [15]. Generally the utilization of AF paint is the most economical and proven manner in most marine structures until present [16]. Besides both hard and ablative antifouling paints, self-polishing copolymer antifouling (SPC-AF) paint is used for the most widely used AF paint due to the consideration of the cost-benefit analysis, and the optimized control of the leaching process for copper compound or cuprous oxide (Cu_2O) as main biocide additives [9]. In addition, due to harmful effects of tributyltin (TBT) as biocides [12], nowadays, tin-free self polishing copolymers (tin-free SPC) antifouling paint utilized as commercial environmentally friendly AF paint, containing copper compounds. Most a generic formulation of that paint consists of resin, pigment, solvent, primary biocide, co-biocide, organic or organo-metallic booster biocides, plasticizer, extender and so on which are constructed as AF paint system.

Furthermore, in last two decades, researchers had investigated extensively the performance of AF paint which related to paint formulations, mechanisms of biofouling inhibition, and the effect of surrounded environment [9, 10, 12], but no or less comprehensive investigation in Indonesia as a representative of tropical country in south east asia region. In addition, the Indonesian researcher has reported the performance of commercial AF paints that exposed in Madura strait, East Java Province [8]. Therefore, the objective of this work is to elucidate the efficacy of perfomance on antifouling paints compared to anticorrosion paint and bare metal as referenced materials in Bali Sea and the dentrimetal effect of corrosion attack on bare metal without protection of AF paint during field exposure in different depth of the sea.

2. MATERIALS AND METHODS

2. 1.The Preparation of Specimens The chemical composition of bare mild steel was 99.0 wt.% Fe, 0.29 wt.% Si, 0.20 wt.% Cr, 0.13 wt% C, 0.11wt% Si, 0.02 wt% P, 0.01 wt% Ni, 0.01 wt% Cu and 0.01 wt.% Ti that used in this work. The measurement of its composition utilized ARL 3460 optical emission spectrometers (OES). The plates of mild steel were made (length: 250 mm, width: 200 mm and thickness: 3 mm) for coating substrate. All steel plates were sanded by sandblasting machine according to ISO 8501-1 Sa 2.5. In addition, that standard specifies that stains, shadows and streaks have to be restricted to 5% of the clean surface steel area in which it applied commonly for off-shore structure, vessel and others. Both commercial AF paints were multilayer coating system which comprises epoxy primer base coat, epoxy intermediate base coat and top coat (SPC-AF Paint). In particular, intermediate coat of Paint A system consist of glass flake as additive, but not Paint B.

Both anticorrosion (AC) paints also were prepared which consist of primer and intermediate coats without addition AF paint. In addition, all anticorrosion paints is a type of epoxy-based paint without addition main biocides. AC paint and bare mild steel were as referenced materials. The bare metal steels were cut into specimens (150 mm x 75 mm x 3 mm) in which were used to measure their corrosion rates through weight loss method. Furthermore, both commercial AF paints were received from two different companies in Indonesia which is based in tin-free self-polishing copolymers (SPC). The formulation process of AF paints was carried out in those companies with following the approximate chemical composition of paints is shown in Table 1.

Each batch of panels consist of AF paint, AC paint and bare metal plates was attached and placed on submerged piles of Mandara Bali Highway during tropical wet season in early September, at depth of 0 to 3

TABLE 1. Antifouling paint composition

Substances	Approx. w/w% antifouling paint A	Approx. w/w% antifouling paint B
Cu_2O	40 - 50	≥25 - ≤50
ZnO	1 - 5	≤5
CuPT	1 - 5	≤5
Xylene	17	≥10 - ≤25
Ethylbenzene	11	<10
Colophony	-	≤5
Rosin	5 - 10	-
Hydrocarbons, C9, aromatics, (<0.1% Benzene)	-	≤5
TiO_2	0.1 - 1	-
Methanol	0.1 - 1	-
Ethanol	0.1 - 1	-
Other substances : Plasticizer, Anti settling agent, Extenders and Anti sagging agents	Remains	Remains

meters in Benoa Bay (8°45'56.9"S 115°11'48.6"E) are shown in Figure 1. After 30 days of exposure, a batch of panels was retrieved from the sea and dry-stored until further characterization and analysis processes. In Figure 2, the test racks which consist of AF paint, AC paint and bare metal plates were embedded on piles of Mandara Bali Highway, Bali Province.

In previous work, the seawater parameters which comprise dissolved oxygen (DO), temperature, pH, salinity and conductivity can affect the corrosion resistance of material [17] and the growth and the settlement of marine biofouling on offshore structure [7]. Therefore, these factors could be considered to be conducted in this recent work. In measuring them, the apparatus in utilized was HACH HQ40d Advanced Portable meter with field probe type.

2. 2. Evaluation of Painting Properties The evaluations of painting properties were carried out before and after field exposure. The measurement of the painting hardness utilized pencil hardness tester Elcometer 501 in which refer to American Standard Testing and Material (ASTM) D-3363. The adhesion strength of coating was conducted by using automatic adhesion tester Elcometer 510 (ASTM D-4541). The gloss property of coating was also done in specified

Figure 1. Location for placing the specimens in Bali Bay, Indonesia

Figure 2. The test rack of specimens for submerged piles of Mandara Bali highway

measurement angle of 60^o using Horiba Gloss Checker IG-331 (ASTM D-523). In our work, in order to make ensure the validity of data, we carried out all measurement of test samples of coating (paint) at least 5 times. In addition, after retrieving paint specimens from the sea, the visual investigation is necessary to observe the appearance and distribution of attached biofouling. The observation of the cross-sectional morphology and elements in both AF paint A and AF paint B was carried out by using Energy Dispersive X-Ray Spectroscopy (EDAX) and JEOL JSM-6390 series Scanning Electron Microscope (SEM).

2. 3. Weight Loss Analysis of Bare Metals After retrieving from the sea, the attached fouling was scraped off from surface of bare mild steel. The corrosion product which was located underneath fouling had been removed by chemical cleaning method according to ASTM G-1. After that, the specimens were rinsed with distilled water, dried with blower, and then weighed to calculate their mass losses. The tests were repeated in twice to make ensure the reliability of the results, the corresponding corrosion rates (CR) were calculated, assuming uniform corrosion over the entire surface of the specimens. The corrosion rate in mils per year (mpy) was calculated from the weight loss using the following formula:

$$CR = \frac{W \text{ x } K}{D \text{ x } A \text{ x } T} \quad (1)$$

where W = weight loss in milligrams, K-factor = a constant (3.45 x 10), D = metal density in g/cm^3, A = area of sample in cm^2, T = time of exposure of the metal sample in hours.

3.RESULTS AND DISCUSSION

3. 1. The Evaluation of Paint Properties and Seawater Parameters after Exposure Figure 3 shows gloss properties of AF paints after exposure in various depth of seawater. The value of AF Paint A and B gloss are 7.0 and 3.0 before exposure, respectively. The magnitude of gloss AF Paint A was higher than that of AF Paint B due to lower solid content of paint A compared to that of paint B [18]. The result of both AF paints was less than 10 Gross Units (GU) in which is categorized as low gloss paint. On the basis of the results obtained, there is no different magnitude of gloss paint in various depth of the sea within a month of exposure. In addition, there is also less difference for the gloss property of both AF paints before or after exposure.

Jaic and Palija [19] had reported that the decrease of the surface roughness of paint could increase gloss level of paint. It implies that there is no alteration of surface roughness for both AF paint A and B during one month after exposure. Moreover, both AF paints are classified

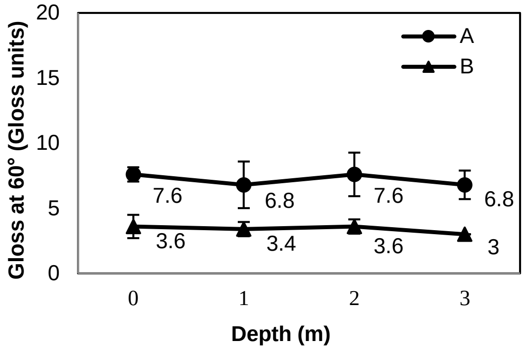
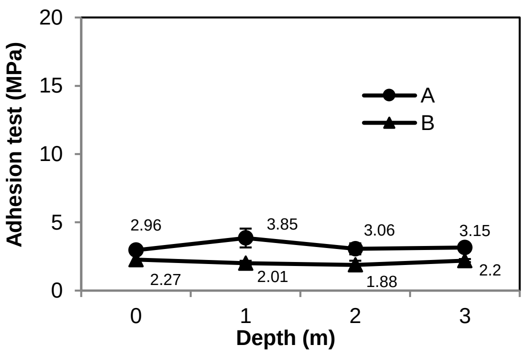


Figure 3. Gloss level of AF Paints as function of depth of the sea

as tin-free self-polishing copolymers (tin-free SPC) in which have self -polishing mechanism for minimizing the level of surface roughness [12]. On the basis of results, the hardness of both AF paint A and Paint B were classified in B before and after exposure in the seawater which are clasified as soft coating. There are no difference of pencil hardness in both AF Paint A and B during exposure

Figure 4a shows adhesion strength of AF Paint A and AF Paint B after exposure in the sea. The adhesion strengths of the paints were almost the same magnitude in different depth level of the sea. The type of failure indicates that the systems of coating tend to 100% cohesion pattern, where the both coating of A and B took place failure inside antifouling coating layer as shown in Figure 4b. In addition, cohesive failure is commonly in the coating itself such as abrasion, cracking due to aging, abrasion, dissolving in solvent and so on.

The habitation of marine fouling organisms is affected by parameters such as salinity, pH, temperature, the intensity of solar radiation and so on [20]. In particular seawater temperature approximately more than 20ºC and pH around 8 are proper for the growth and settlement of marine biofouling [21]. In this work the pH and water temperature tends to become the suitable growth for fouling organisms as shown in Table 2.



(a)

(b)

Figure 4. (a) Adhesion strength for AF paints as function of depth levels of the sea after exposure ; (b) A representative figure of adhesion failure for a spesimen of AF paint after pull off test

TABLE 2. Parameter of seawater after 1 month of exposure in Mandara, Bali Bay

Depth (meters)	Temperature (°C)	pH	Salinity (ppt)	Conductivity (μs/m)	Dissolve Oxygen (mg/L)
0	28.1	8.3	28.9	47.5	5.08
1	28.0	8.3	28.8	47.2	5.19
2	27.9	8.4	28.7	47.1	5.10
3	27.9	8.4	28.6	47.0	5.05

However, there is no significant shift for pH and seawater temperature in various depth of sea level. Furthermore, salinity is defined as the dissolved salt content of water where generally average seawater salinity is 35 ppt [22]. In theory the salinity is assessed by determination of chloride [Cl^-] ion concentrations in the body of water, where empirical dependence, which defines that parameter equals 1.80655 × [Cl^-], is employed. Table 2 shows the magnitude of seawater salinity was less than 35 ppt. Benoa Bay is categorized as a semi-enclosed estuary which has several big river estuaries in Bali Island [23]. It implies that the level of salinity in Benoa Bay is lower than that in ocean outside of coastal area.

In ocean the solubility of dissolved oxygen (DO) decreases as water temperature increases and vice versa where the lower DO concentrations near the equator increase the salinity compared to subtropical area [24]. Commonly oxygen level decrease with increasing depth in the ocean. However, in Table 2, the magnitude of DO of the sea is almost homogenous distributed in different depth. The highest DO concentration took place at the surface layer of the water due to the penetration of intense sunlight and the photosynthesis process in supplying oxygen in the body of water [25]. In this work, the panels only were embedded on piles in shallow depth of the sea up to 3 meters which is in euphotic zone. In that zone the sunlight could penetrate intensively into the body of ocean [26].

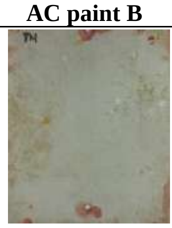
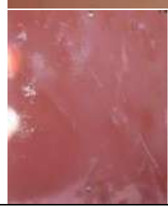
Water conductivity is shifted into salinity by using empirical relationship [17] which is dependent on the concentration of conductive ions present containing the dissolved salt content in the water. According to the results, water conductivity is almost the same magnitude in different depth of the sea as shown in Table 2. The lower water conductivity is indicated to take place due the existence of estuaries near test panels. As reference, generally sea water has a conductivity of about 55 μS/m at temperature of 25°C [27] but not in the present work.

3. 2. The Visual Observation of Paint Specimens after Exposure The visual observation of AC and AF paints was carried out by using digital Canon macro photograph. Figure 5 shows photographs of anticorrosion and antifouling paints A and B before exposure in Benoa Bali Bay. For comparison after field exposure, Table 3 shows the visual observation of all various paints A and B. On the basis of overall visual observation, the utilization of AF paint definitely is paints in different depth of the sea. It is implies that seawater quality parameters is almost the same due to the euphotic zone [26]. In addition, after 1 month of exposure, there were some fouling organisms such as slime, brown

(a) (b)

Figure 5. Photograph of (a) AC and AF paints A and (b) AC and AF paints B before exposure

TABLE 3. The visual observation of specimen after 1 month of exposure

Depth of the sea (m)	AC paint A	AF paint A	Depth of the sea (m)	AC paint B	AF Paint B
0			0		
1			1		
2			2		
3			3		

weed and barnacles on AC paint panels. Barnacles and brown weed are categorized as hard shelled macrofouling and plant macrofouling, respectively [28]. The initial stage of marine fouling growth is commonly found various type bacterial colonies on AF paint which those colonies is as nutrition sources for the growth of invertebrates and microalgae in several days of exposure [29-31]. It is presumed that initial stage was took place in present work. It will be predicted that the growth of biofouling continues to the mature condition over one month of exposure due to suitable environment of seawater.

3. 3. Proposed Mechanism of Antifouling Paint against Marine Biofouling The representative cross section images of both AF paints are shown in Figures 6 and 7. In previous description, both AF paints have three layers coating which consist of primer coating (first layer), intermediate coating (second layer) and top coating (AF paint). The average thickness of both AF paint A and B are 176 µm and 114 µm before exposure, respectively. The reduction of layer thickness in both paints took place after 1 month of field exposure. Furthermore, the AF paint thickness reduction in paint A is lower than that in paint B after exposure as shown in

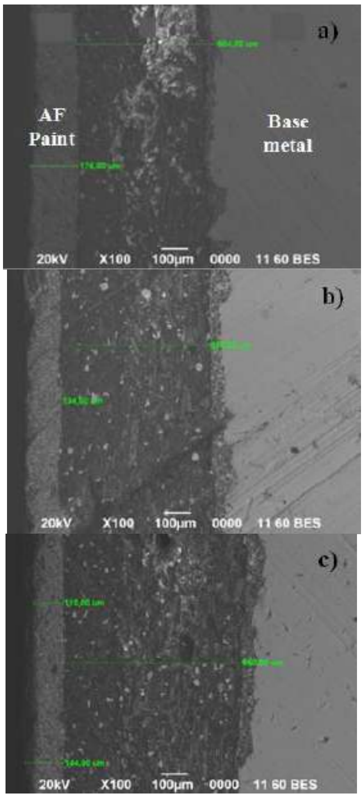


Figure 6. Representative cross-section images of Paint A (a) before exposure and after exposure in seawater at depth of (b) 0 m and (c) 1 m

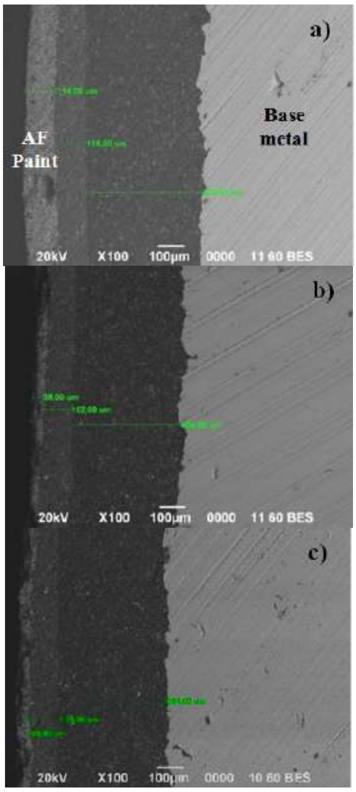


Figure 7. Representative cross-section images of Paint B (a) before exposure and after exposure in seawater at depth of (b) 0 m and (c) 1 m

Figures 6 and 7. The difference of loss thickness of both AF paints is probable due to the difference of their initial thickness. Some papers reported that the service life of AF paint is estimated by the magnitude of initial thickness for coating before exposure [32-34] as well as present study. It implies that the predictive service life of AF Paint A tends to be much longer than that of AF Paint B beyond 1 month of exposure.

In this work, the type of both as-received AF paints is categorized as tin-free-polishing copolymer (tin-free SPC) in which is based on silyl acrylate (SA) polymers as paint binder according to technical data sheet (TDS) of both paints. Tin-free SPC AF paint commonly had been intensively reported [12, 35-37]. The service life of tin-free SPC AF paint commonly is corresponded to the performance of its paint which inhibits the attachment of marine fouling with slow release of binder matrix [12], thus controlling the reduction rate of thickness of the paint.

Furthermore, in the protective mechanism of SPC-AF paint againts the attachment of biofouling, sea water enters into the paint matrix, dissolve such biocides, co-biocides and other additives and diffuse out into the bulk paint again in slow reaction [12]. The thin leached layer of AF paint consist of depleted main biocide and co-biocide particles such Cu_2O and ZnO [12]. In addition, in self-polishing copolymer (SPC) AF paint, the leaching release of paint commonly consist of into initial leaching and steady-state leaching releases [12, 38, 39]. In initial stage, leaching rate starts at once when fresh surface of paint is immersed in seawater and biocide particles begin to dissolve at the interface between the paint layer and water [38, 39]. The absence of soluble pigment biocide will leave behind pores in the matrix paint which is defined as leached layer as show in Figure 8.

AF Paints were formulated by incorporation of binder, solvent, biocide, co-biocide, booster biocide, extender, pigment and other substances as shown in Table 1. In both AF paint A and B, the additive of Cu_2O is as primary biocide in AF paint, where booster biocides such CuPT is used to increase the efficacy of the paint. Co-biocide compound such ZnO is added to improve the performance of the paint by providing enhanced activity against macroalgae fouling and/or algae and bacterial biofilms.

Table 4 shows element composition of AF paint A and AF Paint B before exposure. On the basis of the results, the high concentration of Cu has indication the presence of primary biocides (Cu_2O) and booster biocides (CuPT) in both the AF paints. The element of Zn was also found in both the paints in which refers to the presence of ZnO as co-biocide compound. On the other side, Table 5 also shows element composition of AF paint A and AF Paint B before after exposure. After 1-month exposure, the presence of copper as an essential element of main biocide is still inhibit the settlement of biofouling on the both AF paints. Furthermore, the mechanism of dissolution primary biocide of Cu_2O in seawater in following chemical equation:

$$Cu_2O_{(s)} + 2H^+_{(aq)} + 4Cl_{(aq)} \rightarrow 2CuCl^-_{(aq)} + H_2O_{(l)\,m} \quad (2)$$

The high salinity is induced by the presence of high concentration of chloride ions which increase the dissolution rate of Cu_2O [2]. When cuprous oxide comes into contact with sea water, it generates soluble hydrated Cu(I) chloride complexes which are then rapidly oxidized to Cu^{2+} as the main biocidal species.

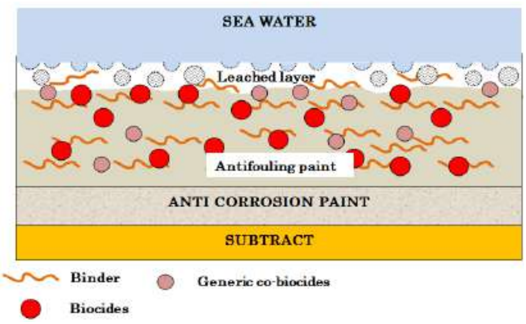


Figure 8. Proposed illustration of nature on antifouling paint system exposed to seawater

TABLE 4. EDAX results of the element composition of AF Paints before exposure

Elements	A	B
	Mass %	
C	36.36	32.42
O	16.44	11.90
Mg	-	1.67
Si	0.73	2.72
S	1.26	1.10
Ti	-	0.81
Fe	-	2.38
Cu	38.34	42.75
Zn	6.23	4.25
Na	-	-
Al	0.62	-
Cl	-	-
Ca	-	-
Ba	-	-

TABLE 5. EDAX results of the element composition of AF Paints after exposure

Elements	A	B
	Mass %	
C	39.38	34.55
O	9.21	12.87
Mg	-	
Si	-	0.85
S	1.26	0.99
Ti	0.64	0.81
Fe	-	-
Cu	27.15	44.24
Zn	2.46	6.50
Na	-	-
Al	-	-
Cl	-	-
Ca	-	-
Ba	-	-

The mechanisms of controlling the release rate of biocides and co-biocides consist of chemical reactions and diffusion where sea water soluble pigment dissolution, binder reaction and paint polishing process occur simultaneously [2]. That mechanism could affect the consistency for thin thickness of leached layer SPC-paint [12, 40]. It implies that the absence of biocides and co-biocides in siylil acrylate matrix leaves behind small pores in that matrix and increases the total wetted area on the paint. The hydrolysis reaction occur throughout the leached layer where there is the shift of wettability of binder from hydrophobic to hydrophilic [12]. Partially reacted binder prone to be eroded by the moving seawater and exposed at a less reacted paint surface (self-polishing effect). The less reacted paint surface consists of biocide and co-biocides enriched matrix which protects further the attachment of biofouling. The condition of steady-state leaching release take places where ion diffusion from seawater through the leached layer is equal to the erosion rate of the paint binder [12].

Furthermore, the release rate of biocide from the paint bulk is induced by current of seawater [41], where leached layer can be easy to polish or erode gradually. Kojima and co-workers [38] reported that the SPC AF paint thickness could be reduced significantly in increasing flow rate of seawater. The slow ocean current occurs in 0.55 m/s which are categorized as ocean drift [42]. Indonesian researchers reported that the range of ocean current around 0.88 m/s to 1.71 m/s where type of that current is affected predominantly by tidal current in Benoa bay [43]. It implies that the ocean current of Bali bay is not slow. The existence of tidal current is able to erode a leached layer of AF paint in which create fresh surface of biocide enriched layer. The fresh layer surface containing biocides is renewed continuously and degradation in paint thickness takes place [44]. The protection of attached biofouling keeps going continuously until there is no AF paint layer on the substrate of steel. The role of pH, salinity and water temperature apparently are not predominant to induce the mechanism of biocide and co-biocide release out from both AF paint matrix due to the same magnitude of the seawater parameters in various depth of the sea. In addition, the geographic location of also lead the different release rate of them in freshwater compared to marine water due to salinity difference [45, 46].

3. 4. Corrosion Behavior of Bare Steel without Coating after Exposure Figure 9 shows the visual observation of bare mild steels before and after field exposure. For quantification of corrosion behavior, Table 6 shows the weight loss and corrosion rate of the bare mild steel after exposure. The weight loss of specimens took place due to the synergism between oxygen reduction in the electrochemical process [17] and the metabolism activity of attached biofouling [4, 8]. The electrochemical reaction is related to the synergic activity between anodic and cathodic area [46-48]. In addition, from the laboratory simulation, the average corrosion rate of steel was 4.1 mpy in seawater which was taken in Mandara Bali without the appearance of fouling on metal substrate [17]. By comparing the present field results, it implies that the loss weight of bare steel apparently increase due to the presence of marine biofouling.

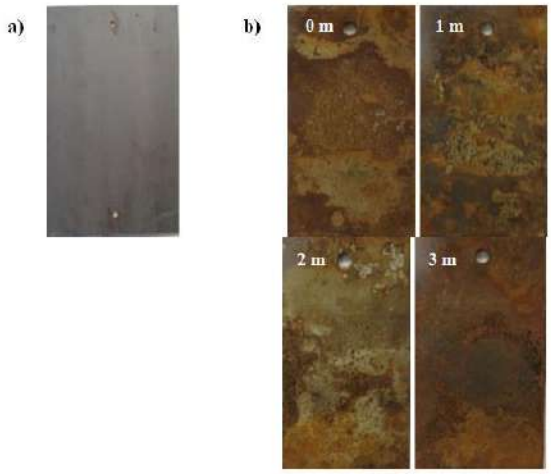


Figure 9. Visual observation of bare mild steel specimens a) before and b) after exposure in different depth of the sea

TABLE 6. The weight loss of specimen in various depth of seawater

Depth of seawater (m)	Weight loss (gram)	Corrosion rate (mpy)
0	6.09	13.03
1	6.075	13.00
2	7.07	15.13
3	6.62	14.17

The magnitude of weight loss of metal is corresponded to corrosion rate as shown in equation 1. On the basis of result, there is no significant difference for the corrosion rate of bare steel at the various depth of seawater. Lutviasari et al. [8] also reported that there is less difference for the magnitude of corrosion rate on bare steel in the seawater depth of 1, 2 and 3 meters in Madura strait, Indonesia as well as the present study. In the case of submerged structures in sea water, the magnitude of corrosion rate is related to the complexity of combination of temperature, salinity, pH, and dissolved oxygen (DO) [47]. In addition, the corrosion rate of bare mild steels took place predominantly due to the effect of DO and biofouling as well as the weight loss of them. Therefore, the presence of antifouling paint with multilayer system decreases the susceptibility of corrosion due to the resistance of corrosion ability in primer coating [48, 49].

4. CONCLUSIONS

The performance of antifouling paint within 1 month of field exposure showed essential results compared to anticorrosion paint and bare mild steel in Bali Sea againts the growth of biofouling. AF paint has remarkable efficacy to protect attached marine fouling organisms, but not AC paint and bare mild steel. The role of various depth of the sea up to 3 meters has no essential effect on the properties of the paints. The presence of copper as a primary biocide could inhibit the growth and habitation of biofouling in both the AF paints. The reduction of thickness for AF paints are predominant to be affected by sea current compared to the parameter of pH, salinity and temperature in different depth of the sea. The magnitude of corrosion rate for bare mild steel is almost same in various depth of the sea where the role of dissolve oxygen and marine bio fouling organisms increase the severity of corrosion during service.

5. ACKNOWLEDGEMENTS

The authors would like to acknowledge the Research Center for Metallurgy and Material – LIPI and Ministry of Research, Technology and Higher Education–Republic of Indonesia for financial support in this research under LIPI Priority and Advanced Research Projects. The authors also thankful to the very helpful support from Research Center and Development of Roads and Bridge, Indonesia's Ministry of Public Works and Housing conducting present research.

6. REFERENCES

1. Peres, R.S., Armelin, E., Moreno-Martínez, J.A., Alemán, C. and Ferreira, C.A., "Transport and antifouling properties of papain-based antifouling coatings", ***Applied Surface Science***, Vol. 341, (2015), 75-85. DOI:10.1016/j.apsusc.2015.03.004
2. Ytreberg, E., Bighiu, M.A., Lundgren, L. and Eklund, B., "Xrf measurements of tin, copper and zinc in antifouling paints coated on leisure boats", ***Environmental Pollution***, Vol. 213, (2016), 594-599. DOI: 10.1016/j.envpol.2016.03.029
3. Hakim, M., Nugroho, B., Nurrohman, M., Suastika, I. and Utama, I., "Investigation of fuel consumption on an operating ship due to biofouling growth and quality of anti-fouling coating", in IOP Conference Series: Earth and Environmental Science, IOP Publishing. Vol. 339, No. 1, (2019), 012037. DOI: 10.1088/1755-1315/339/1/012037
4. Priyotomo, G., "A short review of antifouling paint performance in tropical seawater of indonesia", ***Research & Development in Material Science***, Vol. 8, No. 4, (2018). DOI: 10.31031/rdms.2018.08.000695
5. Schwindt, E., Gappa, J.L., Raffo, M.P., Tatián, M., Bortolus, A., Orensanz, J.M., Alonso, G., Diez, M.E., Doti, B. and Genzano, G., "Marine fouling invasions in ports of patagonia (argentina) with implications for legislation and monitoring programs", ***Marine Environmental Research***, Vol. 99, (2014), 60-68. DOI: 10.1016/j.marenvres.2014.06.006
6. Lord, J.P., Calini, J.M. and Whitlatch, R.B., "Influence of seawater temperature and shipping on the spread and establishment of marine fouling species", ***Marine Biology***, Vol. 162, No. 12, (2015), 2481-2492. DOI:10.1007/s00227-015-2737-2
7. Ferry, M., Wan Nik, W. and Mohd Noor, C.W., "The influence of seawater velocity to the corrosion rate and paint degradation at mild steel plate immersed in sea water", in Applied Mechanics and Materials, Trans Tech Publ. Vol. 554, (2014), 218-221. DOI: 10.4028/www.scientific.net/AMM.554.218
8. Nuraini, L., Prifiharni, S., Priyotomo, G., Sundjono and Gunawan, H., "Evaluation of anticorrosion and antifouling paint performance after exposure under seawater surabaya–madura (suramadu) bridge", in AIP Conference Proceedings, AIP Publishing LLC. Vol. 1823, No. 1, (2017), 020101. DOI: 10.1063/1.4978174
9. Turner, A., "Marine pollution from antifouling paint particles", ***Marine Pollution Bulletin***, Vol. 60, No. 2, (2010), 159-171. DOI: 10.1016/j.marpolbul.2009.12.004
10. Almeida, E., Diamantino, T.C. and de Sousa, O., "Marine paints: The particular case of antifouling paints", ***Progress in Organic Coatings***, Vol. 59, No. 1, (2007), 2-20. DOI: 10.1016/j.porgcoat.2007.01.017
11. Azemar, F., Faÿ, F., Réhel, K. and Linossier, I., "Development of hybrid antifouling paints", ***Progress in Organic Coatings***, Vol. 87, (2015), 10-19. DOI: 10.1016/j.porgcoat.2015.04.007
12. Yebra, D.M., Kiil, S. and Dam-Johansen, K., "Antifouling technology—past, present and future steps towards efficient and environmentally friendly antifouling coatings", ***Progress in Organic Coatings***, Vol. 50, No. 2, (2004), 75-104. DOI: 10.1016/j.porgcoat.2003.06.001
13. Farhat, N., Vrouwenvelder, J.S., Van Loosdrecht, M.C., Bucs, S.S. and Staal, M., "Effect of water temperature on biofouling development in reverse osmosis membrane systems", ***Water Research***, Vol. 103, (2016), 149-159. DOI: 10.1016/j.watres.2016.07.015
14. Hermansyah, H., Atmadipoera, A.S., Prartono, T., Jaya, I. and Syamsudin, F., "Stratification and stability of seawater mass in sulawesi sea", ***International Journal of Sciences: Basic and Applied Research***, Vol. 36, No. 8, (2017), 36-44.
15. Carteau, D., Vallée-Réhel, K., Linossier, I., Quiniou, F., Davy, R., Compère, C., Delbury, M. and Faÿ, F., "Development of environmentally friendly antifouling paints using biodegradable polymer and lower toxic substances", ***Progress in Organic Coatings***, Vol. 77, No. 2, (2014), 485-493. DOI: 10.1016/j.porgcoat.2013.11.012
16. Liu, C., "Development of anti-fouling coating using in marine environment", ***International Journal of Environmental Monitoring and Analysis***, Vol. 3, No. 5, (2015), 373. DOI: 10.11648/j.ijema.20150305.30
17. Sundjono, S., Priyotomo, G., Nuraini, L. and Prifiharni, S., "Corrosion behavior of mild steel in seawater from northern coast of java and southern coast of bali, indonesia, Bandung Institute of Technology, (2017). DOI: 10.5614/j.eng.technol.sci.2017.49.6.5
18. Tiarks, F., Frechen, T., Kirsch, S., Leuninger, J., Melan, M., Pfau, A., Richter, F., Schuler, B. and Zhao, C.-L., "Formulation effects on the distribution of pigment particles in paints", ***Progress in Organic Coatings***, Vol. 48, No. 2-4, (2003), 140-152. DOI: 10.1016/S0300-9440(03)00095-X
19. Jaić, M. and Palija, T., "High gloss finish: The impact of surface roughness on gloss", ***Zaštita Materijala***, Vol. 56, No. 4, (2015), 457-462. DOI :10.5937/zasmat1504457j
20. Bressy, C. and Lejars, M., "Marine fouling: An overview", ***The Journal of Ocean Technology***, Vol. 9, No. 4, (2014), 19-28.
21. Rodriguez, S.G.S., Sithole, N., Dhakal, N., Olive, M., Schippers, J.C. and Kennedy, M.D., "Monitoring particulate fouling of north sea water with sdi and new astm mfi0. 45 test", ***Desalination***, Vol. 454, (2019), 10-19. DOI: 10.1016/j.desal.2018.12.006
22. Suriani, M., Ramlan, S. and Nik, W.W., "Antifouling properties of zinc nitrate in seawater", ***International Journal of Chemical Engineering and Applications***, Vol. 7, No. 5, (2016), 314. DOI: 10.18178/ijcea.2016.7.5.596

23. Rachman, H.A., Hendrawan, I.G. and Putra, I.D.N.N., "Studi transpor sedimen di teluk benoa menggunakan pemodelan numerik", ***Jurnal Kelautan: Indonesian Journal of Marine Science and Technology***, Vol. 9, No. 2, (2016), 144-154. DOI: 10.21107/jk.v9i2.1617

24. Breitburg, D., Levin, L.A., Oschlies, A., Grégoire, M., Chavez, F.P., Conley, D.J., Garçon, V., Gilbert, D., Gutiérrez, D. and Isensee, K., "Declining oxygen in the global ocean and coastal waters", ***Science***, Vol. 359, No. 6371, (2018). DOI : 10.1126/science.aam7240

25. Leidonald, R., Muhtadi, A., Lesmana, I., Harahap, Z. and Rahmadya, A., "Profiles of temperature, salinity, dissolved oxygen, and ph in tidal lakes", in IOP Conference Series: Earth and Environmental Science, IOP Publishing. Vol. 260, No. 1, (2019), 012075. DOI: 10.1088/1755-1315/260/1/012075

26. Jaud, T., Dragon, A.-C., Garcia, J.V. and Guinet, C., "Relationship between chlorophyll a concentration, light attenuation and diving depth of the southern elephant seal mirounga leonina", ***PLoS one***, Vol. 7, No. 10, (2012), e47444. DOI: 10.1371/journal.pone.0047444

27. Sauerheber, R. and Heinz, B., "Temperature effects on conductivity of seawater and physiologic saline, mechanism and significance", ***Chemical Sciences Journal***, Vol. 6, No. 4, (2016), 1-5. DOI: 10.4172/2150-3494.1000109

28. Doble, M., "Polymers in a marine environment, Smithers Rapra, (2014).

29. Lanneluc, I., Langumier, M., Sabot, R., Jeannin, M., Refait, P. and Sablé, S., "On the bacterial communities associated with the corrosion product layer during the early stages of marine corrosion of carbon steel", ***International Biodeterioration & Biodegradation***, Vol. 99, (2015), 55-65. DOI: 10.1016/j.ibiod.2015.01.003

30. Chen, C.-L., Maki, J.S., Rittschof, D. and Teo, S.L.-M., "Early marine bacterial biofilm on a copper-based antifouling paint", ***International Biodeterioration & Biodegradation***, Vol. 83, (2013), 71-76.

31. Mieszkin, S., Callow, M.E. and Callow, J.A., "Interactions between microbial biofilms and marine fouling algae: A mini review", ***Biofouling***, Vol. 29, No. 9, (2013), 1097-1113. DOI: 10.1080/08927014.2013.828712

32. Swain, G.W., Kovach, B., Touzot, A., Casse, F. and Kavanagh, C.J., "Measuring the performance of today's antifouling coatings", ***Journal of Ship Production***, Vol. 23, No. 3, (2007), 164.

33. Atlar, M. and Callow, M., "The development of foul-release coatings for seagoing vessels", ***Journal of Marine Design and Operations***, No. 4, (2003), 11-23.

34. Takahashi, K., Release rate of biocides from antifouling paints, in Ecotoxicology of antifouling biocides. 2009, Springer.3-22. DOI: 10.1007/978-4-431-85709-9_1

35. Lagerström, M., Lindgren, J.F., Holmqvist, A., Dahlström, M. and Ytreberg, E., "In situ release rates of cu and zn from commercial antifouling paints at different salinities", ***Marine Pollution Bulletin***, Vol. 127, No., (2018), 289-296. DOI: 10.1016/j.marpolbul.2017.12.027

36. Amara, I., Miled, W., Slama, R.B. and Ladhari, N., "Antifouling processes and toxicity effects of antifouling paints on marine environment. A review", ***Environmental Toxicology and Pharmacology***, Vol. 57, (2018), 115-130. DOI: 10.1016/j.etap.2017.12.001

37. Chambers, L.D., Stokes, K.R., Walsh, F.C. and Wood, R.J., "Modern approaches to marine antifouling coatings", ***Surface and Coatings Technology***, Vol. 201, No. 6, (2006), 3642-3652. DOI: 10.1016/j.surfcoat.2006.08.129

38. Kojima, R., Shibata, T. and Ueda, K., "Leaching phenomena of antifouling agents from ships' hull paints", ***Journal of Shipping and Ocean Engineering*** , Vol. 6, No. 5, (2016).DOI: 10.17265/2159-5879/2016.05.002

39. Laidlaw, F.B., "The history of the prevention of fouling", ***Marine Fouling and Its Preventionu***, No. 580, (1952), 211–22.

40. Ytreberg, E., Karlsson, J. and Eklund, B., "Comparison of toxicity and release rates of cu and zn from anti-fouling paints leached in natural and artificial brackish seawater", ***Science of the Total Environment***, Vol. 408, No. 12, (2010), 2459-2466. DOI: 10.1016/j.scitotenv.2010.02.036

41. Al Tanto, T., Wisha, U.J., Kusumah, G., Pranowo, W.S., Husrin, S., Ilham, I. and Putra, A., "Karakteristik arus laut perairan teluk benoa–bali", ***Geomatika***, Vol. 23, No. 1, (2017), 37-48. DOI: 10.24895/jig.2017.23-1.631

42. Fay, F., Linossier, I., Peron, J.J., Langlois, V. and Vallee-Rehel, K., "Antifouling activity of marine paints: Study of erosion", ***Progress in Organic Coatings***, Vol. 60, No. 3, (2007), 194-206. DOI: 10.1016/j.porgcoat.2007.07.027

43. Rachmat, B. and Ilahude, D., "Estimation of sea current energy potential by using calculation models of horizontal axis current turbine in toyapakeh strait, nusa penida, bali", ***Bulletin of the Marine Geology***, Vol. 32, No. 2, (2018). DOI: 10.32693/bomg.32.2.2017.352

44. Khavasfar, A., Moayed, M.H. and Jafari, A.H., "An investigation on the performance of an imidazoline based commercial corrosion inhibitor on CO_2 corrosion of mild steel", ***International Journal of Engineering, Transactions A: Basics***, Vol. 20, No. 1, (2007), 35-44.

45. Lagerström, M., Ytreberg, E., Wiklund, A.-K.E. and Granhag, L., "Antifouling paints leach copper in excess–study of metal release rates and efficacy along a salinity gradient", ***Water Research***, Vol. 186, (2020), 116383. DOI: 10.1016/j.watres.2020.116383

46. Wrange, A.-L., Barboza, F.R., Ferreira, J., Eriksson-Wiklund, A.-K., Ytreberg, E., Jonsson, P.R., Watermann, B. and Dahlström, M., "Monitoring biofouling as a management tool for reducing toxic antifouling practices in the baltic sea", ***Journal of Environmental Management***, Vol. 264, (2020), 110447. DOI: 10.1016/j.jenvman.2020.110447

47. Shahrabi, T., Hosseini, M., Ghorbani, M. and Arshadi, M., "Synergistic influence of benzoate ions on inhibition of corrosion of mild steel in 0.5 m sulfuric acid by benzotriazole", ***International Journal of Engineering***, Vol. 16, No. 3, (2003), 255-264.

48. Gu, Y., Yu, L., Mou, J., Wu, D., Xu, M., Zhou, P. and Ren, Y., "Research strategies to develop environmentally friendly marine antifouling coatings", ***Marine Drugs***, Vol. 18, No. 7, (2020), 371. DOI: 10.3390/MD18070371

49. Mahmood, M., Suryanto, S., Al Hazza, M. and Haidera, F., "Developing of corrosion resistance nano copper oxide coating on copper using anodization in oxalate solution", ***International Journal of Engineering***, ***C: Aspects***, Vol. 31, No. 3, (2018), 450-455. DOI: 10.5829/ije.2018.31.03c.07

Persian Abstract

چکیده

برای جلوگیری از رشد بیو فیلم رسوبی دریایی از رنگ های ضد فولیگ استفاده می شود. در اندونزی ، این رنگ به طور گسترده ای برای حمل و نقل استفاده می شود که معمولاً از مواد زیست کش پایه مسی استفاده می شود. در حقیقت ، هیچ مطالعه جامعی در مورد رنگ ضد رطوبت در اندونزی در مقایسه با سایر کشورهای گرمسیری انجام نشده است. در این مطالعه ، ارزیابی عملکرد برای رنگ ضد آلودگی در جایی انجام شد که رنگ ضد خوردگی و فولاد برهنه نیز به عنوان منابع مورد مطالعه قرار گرفتند. اندازه گیری میزان خوردگی بر روی فولاد به روش کاهش وزن انجام شد. پانل های حاوی نمونه ها 1 ماه برای غوطه وری در عمق های مختلف دریا تا 3 متر قرار داشتند. پارامترهای آب دریا متشکل از دما ، pH ، شوری ، رسانایی و اکسیژن محلول و همچنین خصوصیات پوشش اندازه گیری شد. نتایج نشان داد که هر دو سطح رنگ ضد خوردگی و نمونه های فولادی پوشیده شده از بیو فولیگ ، اما نه روی رنگ ضد رسوب گذاری. همچنین در خصوصیات رنگ ضد آلودگی قبل و بعد از قرار گرفتن در عمق های مختلف دریا تفاوت چندانی ندارد. کاهش ضخامت برای رنگ ضد آلودگی ظاهرا غالب است که تحت تأثیر جریان دریا باشد. میزان سرعت خوردگی در فولاد برهنه در عمق مختلف دریا تقریباً یکسان است که به دلیل اثر اکسیژن محلول و سوخت زیستی رخ داده است. در آینده ، مقایسه عملکرد رنگ در تمام مناطق محلی لازم است تا در تمام مناطق محلی اندونزی انجام شود.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 987-998

International Journal of Engineering
Journal Homepage: www.ije.ir

Experimental Study of Polyvinyl Alcohol Nanocomposite Film Reinforced by Cellulose Nanofibers from Agave Cantala

F. Yudhanto[a,b], Jamasri*[a], H. S. B. Rochardjo[a], A. Kusumaatmaja[c]

[a] *Department of Mechanical and Industrial Engineering, Faculty of Engineering, Universitas Gadjah Mada, Indonesia*
[b] *Department of Mechanical Technology, Universitas Muhammadiyah Yogyakarta, Indonesia*
[c] *Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Gadjah Mada, Indonesia*

P A P E R I N F O

Paper history:
Received 20 October 2020
Received in revised form 16 December 2020
Accepted 14 January 2021

Keywords:
Polyvinyl Alcohol
Cellulose Nanofibers
Nanocomposite Film

A B S T R A C T

This paper presents an experimental study of addition of cellulose nanofibers (CNF) extracted by the chemical-ultrasonication process from agave cantala leaf plants in the matrix of polyvinyl alcohol (PVA). Combining these materials produce the nanocomposite film with a thickness of 30 μm. The nanocomposite characteristic was investigated by the addition of CNF (0, 2, 5, 8, and 10 wt%) in PVA suspension (3 wt.%). PVA/CNF nanocomposite films were prepared by a casting solution method. The fibrillation of fibers to CNF was analyzed using Scanning Electron Microscopy and Transmission Electron Microscopy. The nanocomposite film functional group's molecular chemical bond and structural analysis were tested using Fourier Transform Infrared and X-ray diffraction. The PVA/CNF nanocomposite film has significant advantages on the ultraviolet barrier, thermal stability tested by Differential Scanning Calorimetry and Thermogravimetric Analyzer, and tensile strength. Overall, the optimal addition of CNF is 8 wt.% in matrix, resulting in the highest crystallinity index (37.5%), the tensile strength and elongation at break was an increase of 79% and 138%, respectively. It has good absorbing ultraviolet rays (82.4%) and high thermal stability (365ºC).

***doi**: 10.5829/ije.2021.34.04a.25*

1. INTRODUCTION

Cellulose is one of the main components of raw fibers and other constituent materials such as hemicellulose, lignin, and other extractive substances. The cellulose used in this study was obtained from agave cantala leaf plants that grew well in Sumenep regency, Madura, Indonesia. Cellulose was extracted and isolated through several mechanical, chemical, and combined processes to produce the cellulose nanofibers (CNF). The unique structure of the CNF has many advantages, like is biodegradable, biocompatible, large specific surface area, high modulus elasticity, low density, and a low coefficient of thermal expansion [1-4]. Some isolation methods that are often used to obtain CNF are mechanical processes, including the high-pressure homogenizer [5-6], grinding or disc refining [7], cyro-crushing [8], steam explosion [9], and high-speed blender [10].

This process requires a lot of energy and is less efficient in obtaining nanocellulose. Previous studies using isolation by chemical extraction and hydrolysis of strong acids were considered more efficient and did not require high costs and energy to produce nanocellulose. Purified fibers have been cleaned from impurities attached to fibers' surface, such as hemicellulose, lignin, pectin, and other contaminants. The purified fibers were obtained by using chemical extraction processes such as dewaxed, alkalization, and bleaching. The dewaxed process uses a solution of toluene and ethanol. At the same time, alkalization treated uses of potassium hydroxide (KOH), sodium hydroxide (NaOH), and the bleaching treated uses of sodium chlorite ($NaClO_2$), hydrogen peroxide (H_2O_2) solutions has been studied [9, 11-16].

*Corresponding Author Institutional Email: *jamasri@ugm.ac.id* (Jamasri)

Please cite this article as: F. Yudhanto, Jamasri, H. S. B. Rochardjo, A. Kusumaatmaja, Experimental Study of PVA Nanocomposite Film Reinforced by Cellulose Nanofibers from Agave Cantala, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 987-998

CNF isolation carried out using the hydrolysis process such as adding sulphuric acid (H_2SO_4), hydrochloric acid (HCl), nitrit acid (HNO_3) and acetic acid (CH_3COOH). It carried out after the process of purified fibers, and this process aims to open the microcellulose bundle into nanocellulose size. The previous studies have used a chemical extraction method, such as done by Listyanda et al. [17]. The ramie fiber has been isolated by adding sulphuric acid (H_2SO_4) solution of 41, 44, 47, 50 wt.% in the burette tube. The solution of sulphuric acid dropped slowly into the beaker and stirrer acid to pulp ratio (1:20) at the temperature of 45ºC for 30 minutes to produce cellulose nanocrystal (CNC). Additional concentration of H_sSO_4 decrease the crystallinity index (CI) and thermal stability of CNC. The highest crystallinity index was obtained at 41 wt.%, leads from 79.7 to 90.7%. The average diameter and length are 6.26 nm and 106 nm, respectively (aspect ratio is 17).

The reported data by Kusmono and Akbar [18] have isolated the best fiber (ramie) using chemical extraction (alkali-bleaching) and hydrolyzed by hydrochloric acid (HCl). At various concentrations of 6, 8, and 12 M of HCl at the temperature of 45ºC for the different times of hydrolysis were 70, 125, 180 minutes. Additional concentration of HCl does not influence the crystallinity index and thermal stability. The long hydrolysis time was decreased the crystallinity index (CI) and thermal stability. The optimum condition process to produce CNC at a concentration of 6 M of HCl at the temperature of 45ºC for 70 minutes. It creates a rod-like shape of CNC with a high CI from 79.75 to 89.61%, crystallite size 5.81 nm, and the average diameter and length are 8 nm, 158 nm, respectively (aspect ratio is 20).

Krishnadev et al. [19] have isolated the leaf fiber (*Agave Americana*) with chemical extraction (alkali, bleaching) and acid hydrolysis (nitrite acid combined with acetic acid) at temperature 100ºC for 30 minutes. This chemical process, coupled with ultrasonication, produces cellulose nanofibers with average diameters 18.2 ± 10 nm, confirmed by TEM. The CNF has a crystallinity index of 64.1% higher than raw fiber is 50.1%.

Zakuwan and Ahmad [20] have isolated kenaf fiber using 65 wt.% sulphuric acids at a preheat temperature of 45ºC for 40 minutes to produces CNC with average diameter and length of 12-15 nm and 101-260 nm, respectively and result in an aspect ratio of 8-17. The high aspect ratio and crystallinity index were significant in improving the reinforcement of nanocomposites. [21-23].

Nanocellulose is known to their high moisture absorption, which leads to a decrease in mechanical properties. To overcome this problem, adding other hydrophilic biocompatible polymers is necessary to remove the reactivity of hydroxyl groups of the nanocellulose. The polymer used as a matrix in the manufacture of nanocomposite films is polyvinyl alcohol (PVA). PVA is easily dissolved in water and biodegradable polymers, resistance to chemical conditions, so it is an attractive material used as an application based on advanced technology. PVA is also non-toxic on the human body and can be used to manufacture medicine cachets, drug delivery systems, barrier materials, membranes, and yarn for surgery [24-25]. The use of PVA for environmentally friendly food packaging began widely used due to several advantages of PVA, including strength, transparent, lightweight, non-toxic, heat-stable, antimicrobial, and good elasticity [26-28].

There are several parameters characteristics of CNF reinforced PVA nanocomposites such as diameter (D), length (L), aspect ratio (L/D), volume fraction (%V_f), adhesion bond between nanocellulose and matrix, and dispersion the nanocellulose in the matrix [29].

This research's main objective is to produce nanocomposite film, which has high mechanical strength, stiffness, and ultraviolet absorbance by adding gel CNF. The made of CNF by the extraction-ultrasonication process. The mix of CNF and matrix PVA to be homogeneous is an essential factor to obtained the homogeneous solution. Physical properties, UV absorbance, and thermal stability were evaluated by XRD, FTIR, SEM, TEM, UV-Vis (Ultraviolet and Visible) transmittance, as well as DSC and TGA tests. The mechanical properties were tested by tensile strength with UTM (Universal Testing Machine).

2. MATERIALS AND METHODS

2.1. Materials The PVA (C_2H_4O)x (fully hydrolyzed, the density of 1.19 g/cm^3 and molecular weight of 145000 Da) was supplied from Sigma Aldrich. Nanocellulose used was isolated from the Agave Cantala was obtained from Sumenep, Madura island, Indonesia. The chemical solution for the process of dewaxed, alkalization, bleaching, and hydrolysis. The solution used there are toluene, ethanol, sodium hydroxide (NaOH), hydrogen peroxide (H_2O_2) and sulphuric acid (H_2SO_4) were supplied by CV. Wahana Hilab Indonesia, Yogyakarta.

2. 2. Isolation of Cellulose Nanofibers (CNF) Cantala fiber, obtained from the leaf, it was cut into small pieces with a size of approximately 10 mm. Chemical pretreatment in cantala fibers includes dewaxing using a soxhlet contain toluene and ethanol solution (2:1) vol.% for 6 hours (20 cycles) [11], [30]. The alkalization process used 5 wt.% NaOH solution at 100ºC for an hour, followed by bleaching using 3 wt.% H_2O_2 at pH 10 at a temperature of 60°C for an hour.

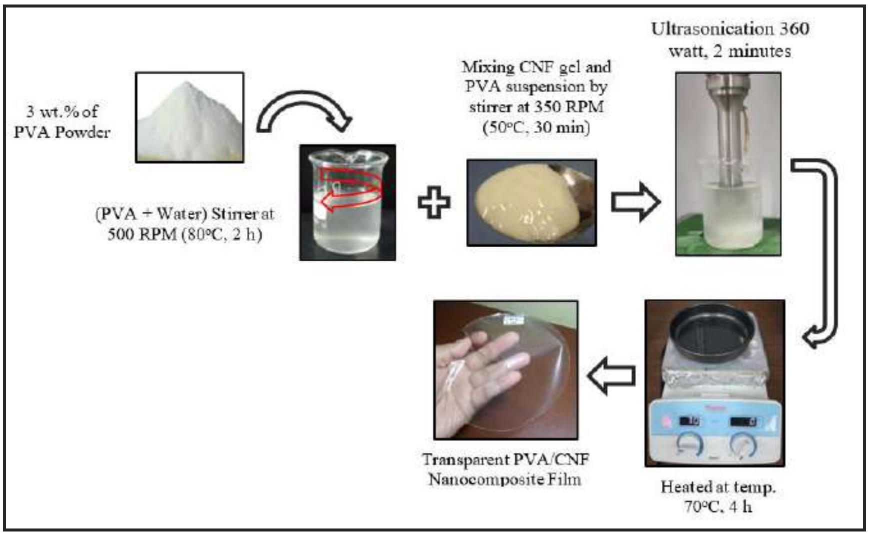


Figure 1. Schemes to produce of PVA/CNF nanocomposites film

The purified fibers were put in an Erlenmeyer flask and placed on a magnetic stirrer with a rotating of 350 RPM [31]. The hydrolysis process used concentration 44 wt.% of the H_2SO_4 solution. It was dripped slowly into an bekker glass from a burette tube. The suspension of CNF was adjusted to constantly rotate slowly with magnetic stirrer (200 RPM) at the pre-heat temperature 60°C for an hour. The suspension of CNF put into the cooled water bath at the temperature of 5°C for 30 minutes to stop acid hydrolysis reaction. Both the suspension was rinsed in the deionized water by centrifuged at 4 cycles and neutralized with NaOH until pH 7.

Fiber fibrillation in the suspension was assisted by an ultrasonic homogenizer (600 watt max. output), which aims to obtain uniform scale of nano sizes. This process was carried out at 20-25 kHz sound waves, 40 % output power (240 watt) using a 6 mm diameter probe for an hour at temperatures of 60°C, and produce a cellulose nanofibers suspension

2. 3. Preparation of PVA/CNF Nanocomposite The CNF liquid suspension was centrifuged at 4000 RPM for 10 minutes to obtain CNF gel. The PVA suspension was prepared by adding the 3 wt.% of PVA powder into pure water at temperature 80°C and stirring at 500 RPM for 2 hours. The PVA suspension was subsequently put into a desiccator containing silica gel for one night to remove the bubbles in the suspension. The CNF gel was prepared with a concentration of 0, 2, 5, 8 and 10 wt.%. The CNF gel and PVA suspension were constantly mixed with stirrer at 350 RPM at temperature 50°C for 30 minutes and followed by ultrasonication with 60% output power (360 watt) for 2 minutes to obtain homogeneous solution [32]. The homogenization process is very important to homogen dispersed CNF in the PVA. The PVA/CNF suspension then casting into a teflon plate, which has dimension is 150 mm in diameter and heated for 4 hours at 70°C (Figure 1).

2.4. Morphology of Fibers Morphology of the fiber was observed by SEM (JSM-6510, LA type JEOL) and TEM (JEM 1400 type, JEOL) characterization to obtain the diameter size (D), and fiber length (L). The Measurement of aspect ratio (L/D) uses imaging analysis program Image J with 20 sample of single nanocellulose. Operating SEM voltage was set 40 kV condition. Test specimen coated with Au and using sputtering technique. Operating TEM voltage was set range from 40 to 120 kV, to get very high contrast and was capable of magnification from 200 to 1,200,000 times.

2. 5. XRD Characterization The XRD patterns of raw fiber and nanocomposite films were measured by using the Rigaku Miniflex-600 type. It was running at 40 kW operating power conditions, and 15 mA current using Cu Kα radiation with wavelength 1.540 Å. Powder sample of XRD test scanned in 2θ range varying from 3° to 40° with an average scan rate of data 2° min^{-1}. The CI (Crystallinity Index) of the organics material is determined by the Segal, as follows in Equation (1) [33]:

$$\mathbf{CI} = \frac{I_{002} - I_{amor}}{I_{002}} \quad \mathbf{(1)}$$

I002 is the intensity of the 002 peaks, which represents the crystalline structure approximately at 2θ = 22°, Iamor is the lowest peak intensity at 2θ=18° represents the amorphous structure.

2. 6. FTIR Characterization FTIR spectra was used to determine the functional bonding groups of materials through changes in the intensity of electromagnetic waves. Wavelengths were measured from a range of 400 to 4000 cm^{-1}. The Shimadzu 8400S spectrometer was used to characterize the materials prepared by blending with potassium bromide (KBr) and followed by pressing to obtain a thin film.

2. 7. Transmittance UV-Vis The ultraviolet (UV) absorption and Visibilty (Vis) of material transparency were characterized by the Ocean Optics USB 4000 spectrophotometer. This characterization was aimed to determine the UV absorption value and its visibility.

2. 8. Mechanical Properties Characterization Tensile strength and elongation at break in the nanocomposite film material were measured by Universal Testing Machine (Pearson Panke Equipment Ltd) with a maximum tensile strength of 200 N and the cross-head speed set of 2 mm/min. The dimension of the tensile test specimen of the plastic film utilized ASTM D-882, for the plastic film thickness from 0.025 to 1 mm, the width and gauge length were 5 mm and 250 mm, respectively.

2. 9. TGA Characterization The thermal stability was performed by a TGA (thermal gravimetry analyzer) and DTG (Diferential Thermal Gravimetry) test was used to calculate the weight of degradation. The TGA characterization is set up in range of 3 to 700 ºC, with a temperature rate of 10 ºC min^{-1} using 50 ml/min nitrogen gas (N_2). The test equipment used is the TGA Mettler Toledo model.

2. 10. DSC Characterization DSC (Differential Scanning Calorimetry) was used to determine thermal parameters and the degree of crystallinity mechanism of materials in a polymer system [34]. The tools used are the DSC-60 Plus and Flow Unit Control, and also use the TA-60WS Collection Monitor software on the desktop. The sample is heated at 30-250 °C with Nitrogen Gas, where the flow rate is 10 ml/min. The degree of crystallinity index could be calculated by estimating enthalpy fusion of endotherms curve of DSC. The degree of CI for polymer as follow the Equation (2) [35],

$$X_c = \frac{\Delta H_f(T_m)}{\Delta H_f{}^o(T_m{}^o)} \quad (2)$$

where Xc is the weight fraction degree of crystallinity, $\Delta H_f\ (T_m)$ is the enthalpy of fusion measured at the melting point T_m and $\Delta H_f^o\ (T_m^o)$ is enthalpy of fusion of totally crystalline the PVA polymer measured at equilibrium melting point (T_m^o) is a 138.6 jg^{-1} [36]. The crystallinity index could be calculated using Equation (3) [35].

$$X_c = \frac{\Delta H_f(T_m)}{\Delta H_f{}^o(T_m{}^o)(1-m_f)} \quad (3)$$

where m_f is the mass fraction filler of CNF in nanocomposite film. The important properties of the polymer system are ussualy related to the structural of crystallinity index (CI).

3. RESULTS AND DISCUSSION

3. 1. Morphology of Cellulose Fibers Figure 2 shows the SEM (Scanning Electron Microscope) image of the cantala fiber before and after the chemical extraction (dewaxed-alkali-bleaching-sulphuric acid hydrolysis 44 wt.%). The bundle raw fibers were composed of individual cellulose linked together by massive cementing material like hemicellulose, lignin and pectin (Figure 2a). It has a diameter between 120-180 μm, the chemical composition of raw agave cantala fiber is cellulose 48.97%, hemicellulose 34.41 %, lignin 11.76 % and other substance 4.86 % [31].

Figure 2b shows the image of treated fiber after dewaxed-alkali process. The most of hemicellulose was damaged and removal became alkali-soluble, but some part of the lignin is still attached to the surface of fibers. The bleaching process aim to remove the lignin structure. The bleaching process with hydrogen peroxide strongly influence the removal most lignin on the surfaces of fibers and also accompanied partially fibrillated the bundle cellulose. The diameter of the cantala fiber after the bleaching process decrease from 30 μm to 60 μm. The surfaces of fiber after the chemical extraction look like clean and more roughness (Figure 2c). In Figure 2d, it is also shown that the diameter of the fiber dramatically changes after the hydrolysis. It causes decreased the diameter of the fiber ranges from 5 ± 2 μm.

Hydrolysis using sulfate acid (H_2SO_4) aims to open a cellulose bundle with a long cellulose chain that will break up into short individual cellulose. The diameter of cellulose bundles also changes from micro to nano size. Sulfate ions have separated the amorphous region (hemicellulose and lignin), and the crystalline region became the individual cellulose nanofibrils, as shown in Figure 3.

The chemical extraction process in fibers produces CPF (chemical purified fiber). It increased the cellulose by 71.58 % and decreased the hemicellulose by 18.45 % and lignin by 6.13 % [31]. The hydrolysis had a significant effect on reduced the lignin structure in the fiber.

The hydrolysis process starts from the breaking of oxygen bonds in the β-1,4-Glycosidic chain, then proceed with the separation of the glycosidic ring bonds and reacts with water (H_2O) were resulted from the

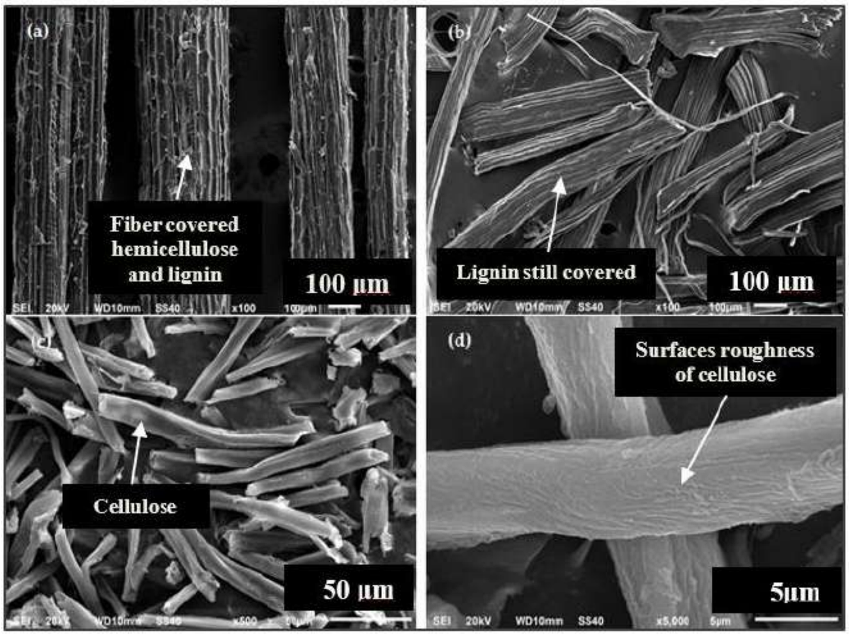


Figure 2. SEM image of (a) raw fiber, (b) treated fiber after alkali, (c) treated fiber after bleaching (d) individual cellulose after sulphuric acid hydrolysis

hydrolysis process of strong acids. Finally, the opening phase of the cellulose rings. Sulfate Acid is a strong acid, in water (H_2O) fully ionized to form hydronium ions (H_3O^+) and hydrogen sulfate (HSO_4^{2-}) (Figure 4).

Ultrasonication process after the hydrolysis process results in an individual CNF which bond together forming the web-like structure could be shown in the TEM image (Figure 5). The web-like structure shaped of cellulose nanofibers (CNF) were calculated by image-J of their length (L), diameter (D) to obtain the aspect ratio (L/D). It is well known that the aspect ratio of CNF affected the crucial role in their reinforcing capabilities [15,31]. The average diameter and length of CNF agave cantala are 45 nm and 1975 nm, respectively; resulting the aspect ratio is 43,8 (Figure 5), it show the high aspect ratio of CNF.

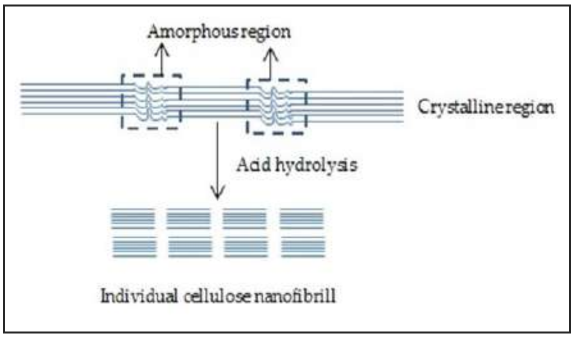


Figure 3. Schematic illustration disintegrated of amorphous region and crystalline region

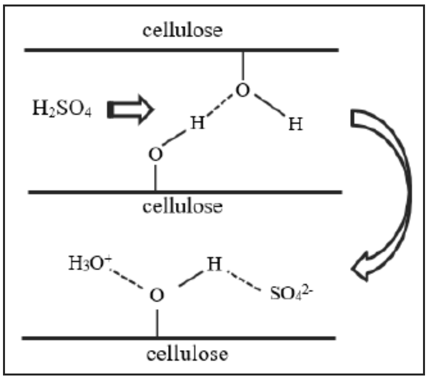


Figure 4. Schematic illustration of chemical reaction of bundle cellulose to individual cellulose by sulphuric acid hydrolysis

The previous study by Siqueira et al. [42], a similar *Agavaceae* family species, the aspect ratio of CNF agave sisalana fibers is 43. The aspect ratio from another fibers that were extracted from agronomy plant source such as coconut husk by 5.5 ± 1.5 nm [37], rice straw by 17 nm, poplar wood by 43 nm, white straw by 45.2 nm [38], soy hulls by 4.43 ± 1.20 nm [39], rice straw by 17 nm [40] and sugarcane bagasse by 4 ± 2 nm [41].

3. 2. Morphology of Nanocomposite Film
Figure 6 is SEM image shows the neat PVA nanocomposite films without the CNF. The neat PVA nanocomposite films show the transparent film, and the

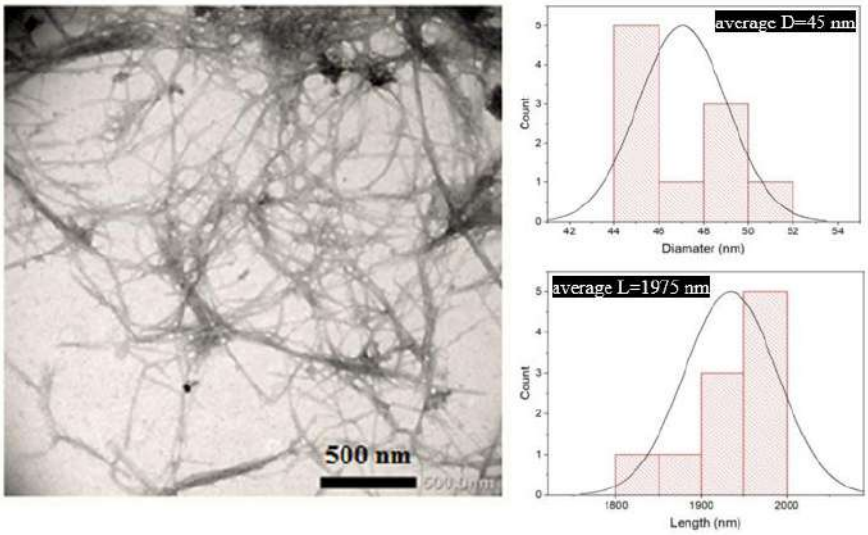


Figure 5. TEM image of diameter and length distribution of cellulose nanofibers (CNF)

polymers distributed uniformly. Figure 7 shows of the 8 wt.% of CNF in the PVA matrix, and it appears that CNF (web like structure) could be well distributed. The good dispersion of the CNF is the major factor to determine the mechanical properties of the nanocomposite film.

However, we founded the agglomerate CNF in the small areas in PVA nanocomposite film. Addition of 10 wt.% CNF causes the aglomeration evenly distributed in some areas on PVA matrix lead to 10-12% (Figure 8). The concentration of 10wt.%. of CNF causes the occurrence of piles of nanocellulose which formed the agglomeration.

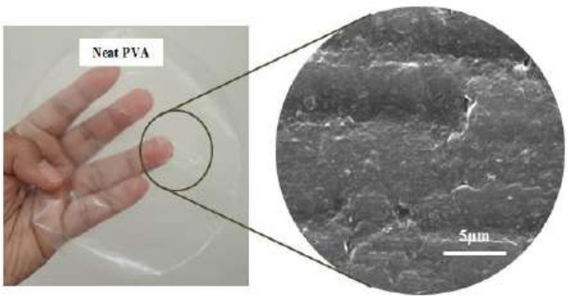


Figure 6. Photo SEM of Neat PVA nanocomposite film

3. 3. XRD Characterization Nanocomposite Film

Figure 9 (a) shows the X-ray diffractogram graph of raw fiber and CNF. The major intensity patterns in the CI of the raw fiber and CNF are peaks at $2\theta = 16.49^o$, 22.84^o and 34.88^o, from the JCPDS (Joint Committee on Powder Diffraction Standards) card of native cellulose (PDF # 030289) shows the crystalline plane is 111, 002, and 040, respectively; indicates the cellulose type I [31, 43, 50]. The amorphous region at the peak a $2\theta = 18^o$. The highest intensity peaks at a $2\theta = 22.84^o$ show a significant increase of cellulose, indicating the amorphous material such as hemicellulose and lignin removal. The correlation changes of the crystalline plane structure from alpha-cellulose ($I\alpha$) to beta cellulose ($I\beta$) due to the chemical-ultrasonication process increase the CI from 64.5% to 78.2%. The beta

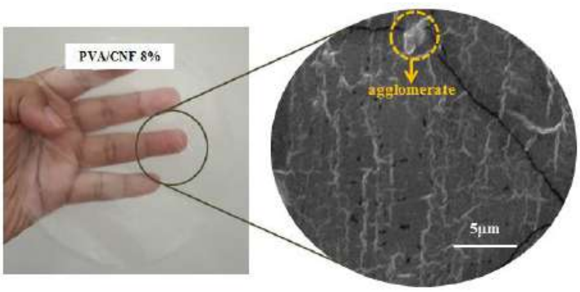


Figure 7. Photo SEM of PVA/CNF 8% nanocomposite film

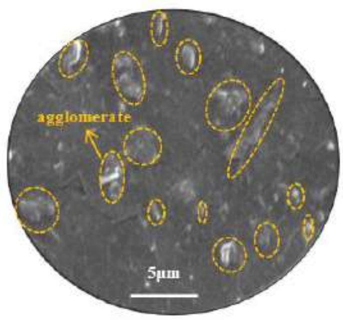


Figure 8. Photo SEM of PVA/CNF 10% nanocomposite film

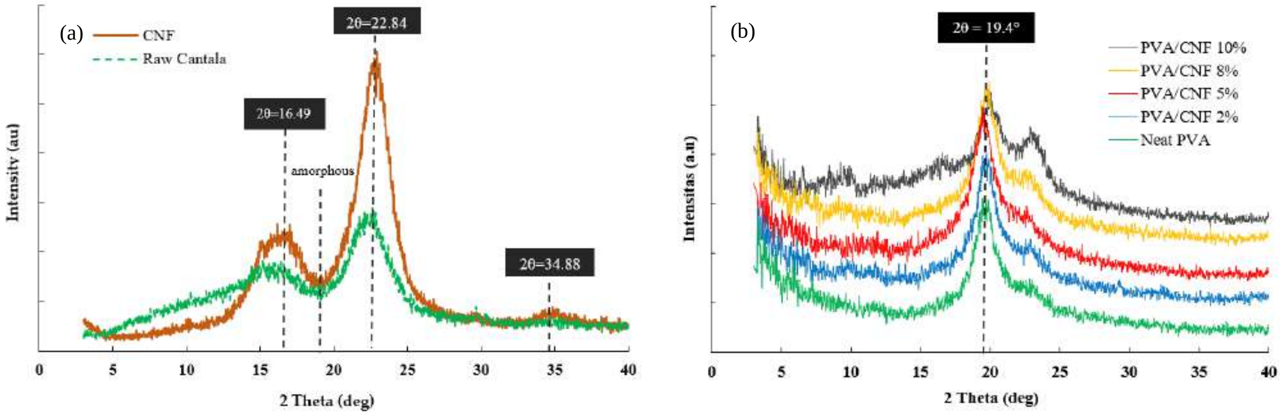


Figure 9. The XRD patterns of (a) raw fiber and CNF, (b) various nanocomposite film

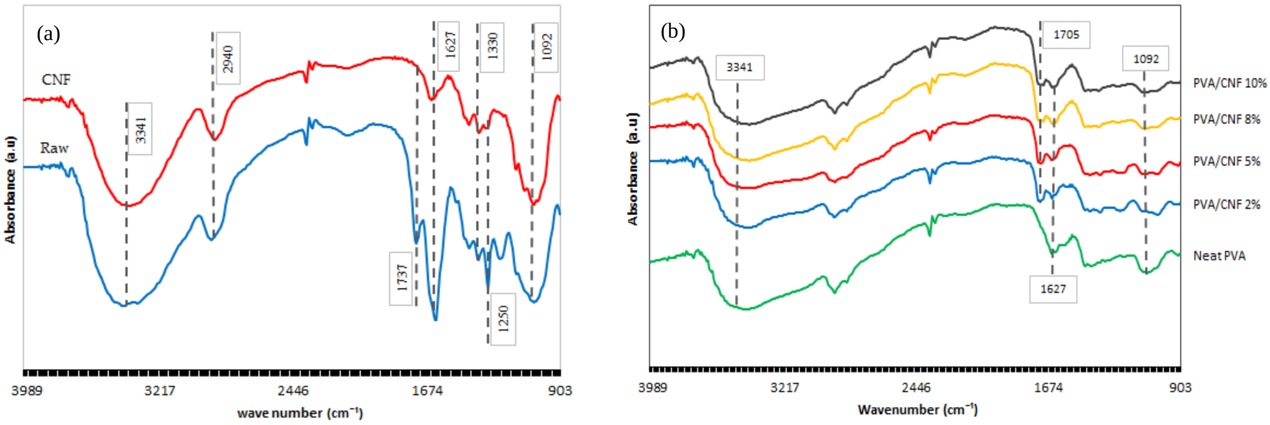


Figure 10. The FTIR patterns of (a) raw fiber and CNF, (b) various nanocomposite film

cellulose (Iβ) has two cellulose hydrogen bonds and is more stable, stronger than alpha cellulose which has only one cellulose hydrogen bond [32, 41].

The X-ray diffractions nanocomposite films showed in Figure 9 (b). The crystallinity index of neat PVA film at 2θ=19.4 is 30.7% compared to the other nanocomposite film. The addition CNF of 2, 5, and 8 wt.% had increased the CI values to 32.0%, 35.8%, and 37.5%, respectively. The relation between increasing the CI values and stiffness of cellulose, where the high CI could be raised the stiffness of cellulose [43]. The higher mechanical strength of nanocomposite film correlated with a higher crystallinity index of CNF. More than 8% CNF in the matrix caused a decrease of CI by 22.9% and impacted mechanical properties decreased.

3. 4. FTIR Characterization of Nanocomposite Film

Figure 10 (a) shows the FTIR spectroscopy to determine the physical structure and functional groups of the lignocellulosic plant. The absorbance peak of a raw fiber at 1250 cm^{-1} corresponds to C-O stretching in the acetyl group. The C=O stretching on carbonyl in ester bonds is shown at wavenumber 1737 cm^{-1} [38, 43-44] there are indicating the presence of hemicellulose and lignin.

Many functional groups of acetyl and uronic ester groups forming hemicellulose and carboxylic groups of ferulic lignin-forming [38]. The peak wavelength of 1627 cm^{-1} indicates the absorbance of water [12-13], [15].

The absorption 1092 cm^{-1} the asymetry ring pyranose C-O-C showed that existance the content of an hydroglucose in the cellulose I [45]. The absorption at 1330 cm^{-1} is attributed to CH_2 symmetric bending and C-O aromatic ring polysaccharides. The peak of 2924 cm^{-1} shows aliphatic saturated C-H stretching vibration in cellulose [15]. The wavenumber 3340-3342 cm^{-1} shows the intramolecular hydrogen bonding (-OH) group of cellulose [45].

Figure 10 (b) shows the FTIR spectra graph of PVA nanocomposite films' peak in various addition of CNF.

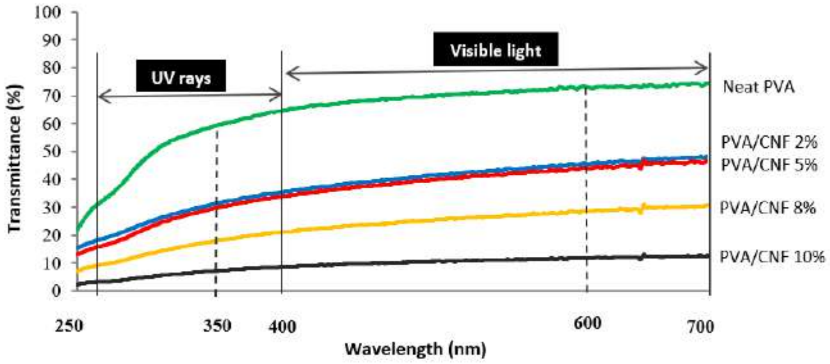


Figure 11. The UV-Vis spectra anaysis of nanocomposite film

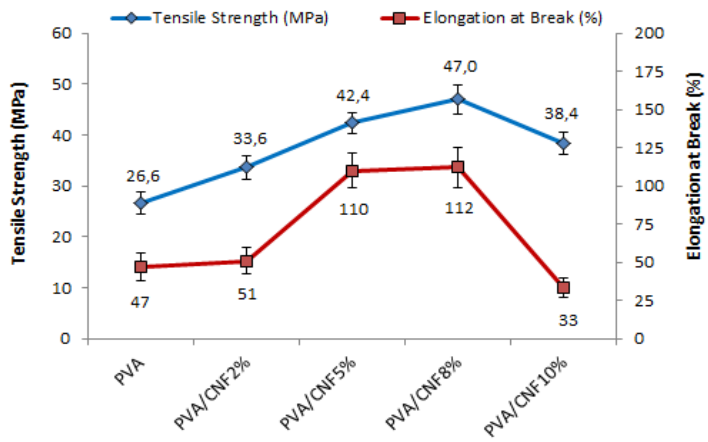


Figure 12. The Tensile strength and elongation at break of nanocomposite film

The sharp peaks waveform change into ramps at the range of 1092 cm^{-1} (C-O-C stretching) and 3341 cm^{-1} shows that monosaccharides that cellulose bond-forming begin to strongly cross-linked with the matrix. In the previous research conducted by Lim et al. [46], PAA hydrogel material and cellulose nanocrystals showed a good bond even though the two materials retain their unique characteristics which changes–the intensity and shape of peaks at waveform 1054-1092 cm^{-1}. Addition of CNF also reduces the water absorbance, which may cause the adhesion bonding between hydrophilic properties in PVA with CNF material, as shown waveform pattern at 1627cm^{-1}.

3. 5. The UV-Vis Absorbance of Nanocomposite Film Figure 11 shows that the CNF addition is also affecting the transparency of the film, which seems to decrease slightly. The visible light spectroscopy, it is shown that the neat PVA nanocomposite film has an excellent visible light transmittance is 73%. Addition of 2, 5, 8, and 10 wt.% of the CNF in the matrix had decreased the visible light transmittance to 45, 43.3, 28 and 11.7%, respectively, but raises the absorbance of The UV rays at 350 nm wavelength.

The percentage absorb of UV rays in the PVA nanocomposite is a little (41.8%). Addition of 2, 5, 8, and 10 wt.% of CNF causes an increase in the absorb UV rays significantly became 69.2, 70.8, 82.4, 92.9 %. The small transmittance is indicating the material has strong absorbance of UV rays. The cellulose and lignin in the CNF significantly influence the absorption of UV rays and visible light. The lignin has a complex structure polydispersity in molecular weight to absorb the UV [47]. The excellent UV absorbance and still good transparency are PVA/CNF 8% (UV absorbance is 82.4%, and visible transmittance light is 28%).

3. 6. Mechanical Properties Figure 12 indicates the effect of CNF on tensile strength and elongation at the break of the nanocomposite films. From the mechanical test, addition of CNF to the PVA matrix would increase the nanocomposite's mechanical properties. Generally, an increase in CNF would increase the tensile strength and elongation at break.

The neat PVA has 26.6 MPa in tensile strength and 47% in elongation at break. Addition of 2 wt.% CNF did not give significant effect at mechanical properties because the CNF undispersed in the PVA matrix. Whereas the addition of 8 wt.% of CNF gives a substantial improvement in the mechanical properties, It increase the tensile strength and elongation at break are 79% and 138%, respectively. This condition shows that CNF has distributed well in the matrix, and it causing the homogeneous stress of nanocomposite film. The homogeneous stress leads that the excellent polymer branching bonds between the hydroxyl group (-OH) in the nanocellulose and carbonyl groups (C=O) in the matrix's polymer chain.

3. 7. DSC Thermal Analysis DSC (Differential Scanning Calorimetry) is one of the necessary tests to determine the thermal parameter such as a glass transtition (T_g), phase changes, melting purity crystallization (T_m), heat capacity (ΔH_m) and degree of crystallinity (X_c) of the polymer. According to Agrawal et al. [48] the PVA is one of the partially crystalline polymers exhibiting both the glass transition temperature, T_g (characteristic of amorphous phase) and melting iso-therm, T_m (characteristic of crystalline phase).

Figure 13 and Table 1 show the glass transition and melting temperature on Neat PVA and PVA/CNF nanocomposite films. The neat PVA films show glass transition temperature at 85ºC. The addition of CNF 2, 5 and 8 wt.% increases the glass transition and the melting temperature. Addition of CNF 10 wt.% causes the decrease the glass transition and melting temperature. It's causes chemical bonding starting weakness and then the hydroxyl group in cellulose easily absorbs the heat.

This resutl is slighty different with the previous research by Patel et al. [35], addition of 5, 10, 15, and 20 wt.% of palm leaf (PL) nanocrystalline powder in PVA films has increased melting temperature (T_m), but it did not happen on a glass transition temperature (T_g). That addition does not significant change T_g material of PL/PVA.

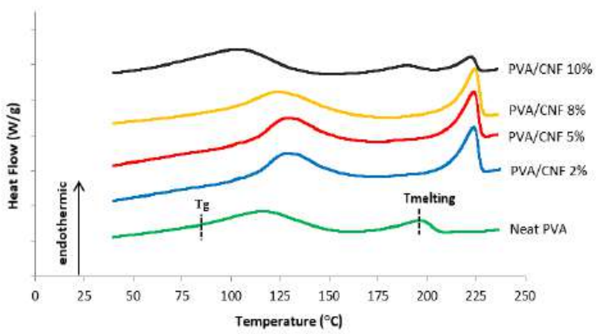


Figure 13. The DSC melting result of nanocomposite film

TABLE 1. The transition glass and melting temperature nanocomposite film

Samples	T_g (in ºC)	T_m (in ºC)	ΔH_m (in J/g)	X_c (in %)
Neat PVA	85	195	30.8	22.2
PVA/CNF 2%	106	230	48.6	35.8
PVA/CNF 5%	108	225	47.3	35.9
PVA/CNF 8%	102	225	55.5	43.5
PVA/CNF 10%	65	189	20.2	16.2

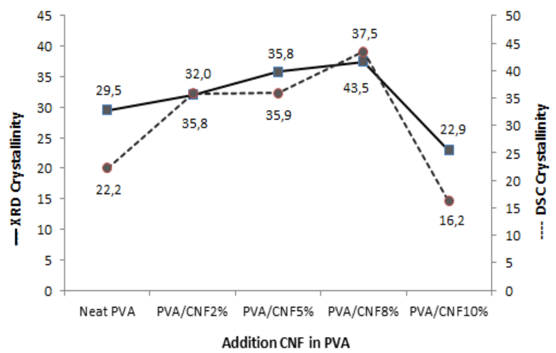


Figure 14. The crystallinity index from XRD and DSC test

Figure 14 shows the enthalpy fusion of DSC could determine the degree of crystallinity index (CI). The CI pattern from DSC results has a similar correlation with XRD spectra result. Addition of CNF from 2 to 8 wt.% had increased the CI of nanocomposite film and after more than 8 wt.%, the CI decreased. It's indicated the some crystalline region starting degraded in the molecular bond at initial temperature.

3. 8. Thermogravimetric Analysis Figures 15 and 16 show the thermogravimetric Analysis (TGA) and derivative thermogravimetric (DTG) of nanocomposite film. The weight loss and derivative weight of nanocomposite films that occur by changing the temperature, which indicated by T_{onset} (the initial degradation of material), $T_{10\%}$ (the weight loss of 10% material), and T_{max} (the maximum of degradation). In this study the T_{onset} to T_{max} of neat PVA film occur at temperature from 213 ºC to 275 ºC (MW = 145,000 Da), slightly higher than the the previous study by Rynkowska et al. [49] at temperature from 210 ºC to 265 ºC (MW = 100,000 Da), its due to the differences in molecular weight (MW).

Typically, initial decomposition of the lignocellulosic material has been started above at 200ºC, it occurs due to the evaporation of adsorbed moisture and the remaining hemicellulose's burning earlier. In all cases, the initial weight loss quickly occurs in the remaining isolation results in the hemicellulose contain the acetyl functional groups. The decomposition of the hemicellulose region firstly occurs before lignin and cellulose [37]. Addition of CNF as reinforcement in PVA can increases the thermal stability of nanocomposite films. This behavior caused the CNF has a well alignment crystallite structure so that the nanocomposite film has high stability thermal at the temperature above at 300ºC.

Table 2 shows the increase of thermal stability in each variation of nanocomposite film. Addition of CNF 2 to 5 wt.% could increase maximum thermal degradation (T_{max}) to 310ºC, 340ºC and 365ºC, respectively. Addition of 10 wt.% CNF causes the slightly decreases the thermal stability to 363ºC. It's

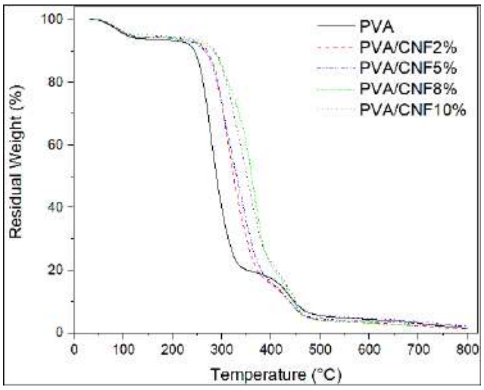


Figure 15. The graph of Thermal Gravimetri Analyze (TGA)

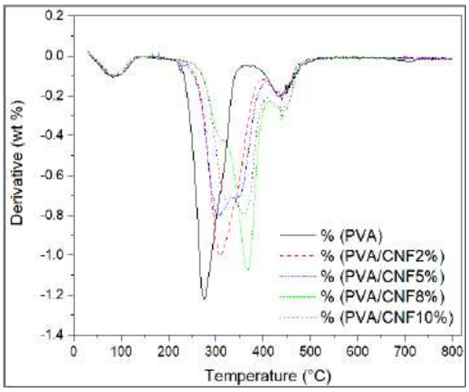


Figure 16. The graph of Differential Thermal Gravimetry (DTG)

TABLE 2. TGA Result of Neat PVA and PVA/CNF nanocomposite film

Samples	T_{onset} (in ºC)	T_{max} (in ºC)	$T_{10\%}$ (in ºC)
PVA	213	275	244
PVA/CNF 2%	225	310	270
PVA/CNF 5%	215	340	265
PVA/CNF 8%	251	365	282
PVA/CNF 10%	250	363	285

caused the crystallite structure linkages between cellulose and the matrix beginning unstable.

4. CONCLUSION

The effect of 8 wt.% CNF on the PVA film matrix can improve the physical properties and mechanical strength. The FTIR and XRD test shows that a good dispersion and high mechanical interlocking between CNF and PVA. The crystallinity index was increased from 29.5% to 37.5%. The tensile strength and elongation at break were increased from 26.6 to 47 MPa and 47% to 112%, respectively. The high thermal stability indicates the shift lead the glass transition temperature (T_g) from 85ºC to 102ºC and the maximum degradation temperature (T_{max}) from 275ºC to 365ºC. Based on the characterization results, it's shown that nanocomposite film has the highest mechanical strength, UV barrier, transparent, and high thermal stability.

5. ACKNOWLEDGEMENT

This research was carried out under a dissertation grant by Ministry of Finance of The Republic Indonesia through Lembaga Pengelola Dana Pendidikan (LPDP) BUDI-DN scholarship program No. PRJ-1427/LPDP.4/2019.

6. REFERENCES

1. Kumar, Ritesh, Sanju Kumari, Shivani Singh Surah, Bhuvneshwar Rai, Rakesh Kumar, Sidhharth Sirohi, and Gulshan Kumar. "A simple approach for the isolation of cellulose nanofibers from banana fibers." ***Materials Research Express,*** Vol. 6, No. 10, (2019), 105601. doi: 10.1088/2053-1591/ab3511
2. Syafri, Edi, Anwar Kasim, Hairul Abral, and Alfi Asben. "Cellulose nanofibers isolation and characterization from ramie using a chemical-ultrasonic treatment." ***Journal of Natural Fibers,*** (2018). doi: 10.1080/15440478.2018.1455073
3. Meng, Fanrong, Guoqing Wang, Xueyu Du, Zhifen Wang, Shuying Xu, and Yucang Zhang. "Extraction and characterization of cellulose nanofibers and nanocrystals from liquefied banana pseudo-stem residue." ***Composites Part B: Engineering,*** Vol. 160, (2019), 341-347. doi: 10.1016/j.compositesb.2018.08.048
4. Moon, Robert J., Ashlie Martini, John Nairn, John Simonsen, and Jeff Youngblood, "Cellulose nanomaterials review: structure, properties and nanocomposites." ***Chemical Society Reviews***, Vol. 40, (2011), 3941-3994. doi: 10.1039/c0cs00108b
5. Panyasiri, Panee, Naiyasit Yingkamhaeng, Nga Tien Lam, and Prakit Sukyai. "Extraction of cellulose nanofibrils from amylase-treated cassava bagasse using high-pressure homogenization." ***Cellulose,*** Vol. 25, No. 3, (2018), 1757-1768. doi: 10.1007/s10570-018-1686-6
6. Ilyas, R. A., S. M. Sapuan, M. R. Ishak, and E. S. Zainudin. "Sugar palm nanofibrillated cellulose (Arenga pinnata (Wurmb.)

Merr): effect of cycles on their yield, physic-chemical, morphological and thermal behavior." ***International Journal of Biological Macromolecules***, Vol. 123 (2019), 379-388. doi: 10.1016/j.ijbiomac.2018.11.124

7. Supian, Muhammad Arif Fahmi, Khairatun Najwa Mohd Amin, Saidatul Shima Jamari, and Shahril Mohamad. "Production of cellulose nanofiber (CNF) from empty fruit bunch (EFB) via mechanical method." ***Journal of Environmental Chemical Engineering,*** Vol. 8, No. 1, (2020), 103024. doi: 10.1016/j.jece.2019.103024

8. Bhatnagar, A., and M. Sain. "Processing of cellulose nanofiber-reinforced composites." ***Journal of Reinforced Plastics and Composites,*** Vol. 24, No. 12, (2005), 1259-1268. doi: 10.1177/0731684405049864

9. Song, Yan, Wei Jiang, Yuanming Zhang, Haoxi Ben, Guangting Han, and Arthur J. Ragauskas. "Isolation and characterization of cellulosic fibers from kenaf bast using steam explosion and Fenton oxidation treatment." ***Cellulose,*** Vol. 25, No. 9, (2018), 4979-4992, doi: 10.1186/s40201-015-0167-1

10. Rochardjo, H.S.B Jamasri and Yudhanto, F. "Extraction of Natural Fibers by High-Speed Blender to Produce Cellulose Sheet Composite." ***International Review of Mechanical Engineering,*** Vol. 13, No. 12, (2019), 691-699. doi: 10.15866/ireme.v13i12.17586

11. Abe, Kentaro, and Hiroyuki YaNo. "Comparison of the characteristics of cellulose microfibril aggregates of wood, rice straw and potato tuber." ***Cellulose,*** Vol. 16, No. 6, (2009), 1017-1023. doi: 10.1007/s10570-009-9334-9.

12. Yudhanto, F. and Rochardjo, H.S., "Application of taguchi method for selection parameter bleaching treatments against mechanical and physical properties of agave cantala fiber." In ***IOP Conference Series: Materials Science and Engineering***, IOP Publishing, Vol. 352, (2018), 012002. doi:10.1088/1757-899X/352/1/012002

13. Yudha, V., Rochardjo, H.S.B., Jamasri, J., Widyorini, R., Yudhanto, F. and Darmanto, S., "Isolation of cellulose from salacca midrib fibers by chemical treatments." In ***IOP Conference Series: Materials Science and Engineering***, IOP Publishing, Vol. 434, (2018**)**, 012078. doi: 10.1088/1757-899X/434/1/012078

14. Carvalho, Kelly Cristina Coelho, Daniella Regina Mulinari, Herman Jacobus Cornellis Voorwald, and Maria Odila Hilário Cioffi, "Chemical modification effect on the mechanical properties of hips/coconut fiber composites." ***BioResources***, Vol. 5, (2010**)**, 1143-1155. doi: 10.4186/ej.2012.16.2.73

15. Kargarzadeh, Hanieh, Ishak Ahmad, Ibrahim Abdullah, Alain Dufresne, Siti Yasmine Zainudin, and Rasha M. Sheltami, "Effects of hydrolysis conditions on the morphology, crystallinity, and thermal stability of cellulose nanoc extracted from kenaf bast fibers." ***Cellulose***, Vol. 19, (2012), 855-866. doi: 10.1007/s10570-012-9684-6

16. Rizal, Samsul, Deepu Gopakumar, Sulaiman Thalib, Syifaul Huzni, and H. Abdul Khalil, "Interfacial compatibility evaluation on the fiber treatment in the typha fiber reinforced epoxy composites and their effect on the chemical and mechanical properties." ***Polymers,*** Vol. 10, (2018), 1316. doi: 10.3390/polym10121316.

17. Listyanda, R. Faiz, Kusmono, Muhammad Waziz Wildan, and Mochammad Noer Ilman. "Extraction and characterization of nanocrystalline cellulose (NCC) from ramie fiber by sulphuric acid hydrolysis." In ***AIP Conference Proceedings***, Vol. 2217, No. 1, p. 030069. AIP Publishing LLC, (2020). doi: 10.1063/5.0001068.

18. Kusmono, K. and Akbar, D.A., "Influence of Hydrolysis Conditions on Characteristics of Nanocrystalline Cellulose Extracted from Ramie Fibers by Hydrochloric Acid Hydrolysis." (2020), doi: 10.21203/rs.3.rs-31322/v1.

19. Krishnadev, P., Subramanian, K.S., Janavi, G.J., Ganapathy, S. and Lakshmanan, A., "Synthesis and Characterization of Nano-fibrillated Cellulose Derived from Green Agave americana L. Fiber." ***BioResources,*** Vol. 15, No. 2 (2020) 2442-2458. doi: 10.15376/biores.15.2.2442-2458

20. Zarina, Siti, and Ishak Ahmad. "Biodegradable composite films based on κ-carrageenan reinforced by cellulose nanocrystal from kenaf fibers." ***BioResources,*** Vol. 10, No. 1 (2015), 256-271. doi: 10.15376/biores.10.1.256-271

21. Kargarzadeh, Hanieh, Marcos Mariano, Jin Huang, Ning Lin, Ishak Ahmad, Alain Dufresne, and Sabu Thomas, "Recent developments on nanocellulose reinforced polymer nanocomposites: A review." ***Polymer,*** Vol. 132, (2017) 368-393. doi: 10.1016/j.polymer.2017.09.043

22. Teja Prathipati, S.R.R., Rao, C.B.K. and Dakshina Murthy, N.R., "Mechanical behavior of hybrid fiber reinforced high strength concrete with graded fibers." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 33, No. 8 (2020), 1465-1471. doi: 10.5829/IJE.2020.33.08B.04

23. Chen, Wenshuai, Kentaro Abe, Kojiro Uetani, Haipeng Yu, Yixing Liu, and Hiroyuki Yano, "Individual cotton cellulose nanofibers: pretreatment and fibrillation technique." ***Cellulose,*** Vol. 21, No. 3, (2014), 1517-1528. doi: 10.1007/s10570-014-0172-z

24. Zhang, Wei, Zhennan Zhang, and Xinping Wang, "Investigation on surface molecular conformations and pervaporation performance of the poly(vinyl alcohol) (PVA) membrane." ***Journal of Colloid and Interface Science,*** Vol. 333, (2009), 346-353. doi: 10.1016/j.jcis.2009.01.058

25. Tang, Yufeng, Yumin Du, Yan Li, Xiaoying Wang, and Xianwen Hu, "A thermosensitive chitosan/poly (vinyl alcohol) hydrogel containing hydroxyapatite for protein delivery." ***Journal of Biomedical Materials Research Part A***: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials, Vol. 91, (2009), 953-963. doi: 10.1002/jbm.a.32240

26. Ghaderi, Moein, Mohammad Mousavi, Hossein Yousefi, and Mohsen Labbafi, "All-cellulose nanocomposite film made from bagasse cellulose nanofibers for food packaging application." ***Carbohydrate Polymers,*** Vol. 104, (2014), 59-65. doi: 10.1016/j.carbpol.2014.01.013

27. Perumal, Anand Babu, Periyar Selvam Sellamuthu, Reshma B. Nambiar, and Emmanuel Rotimi Sadiku, "Development of polyvinyl alcohol/chitosan bio-nanocomposite films reinforced with cellulose nanocrystals isolated from rice straw." ***Applied Surface Science,*** Vol. 449, (2018), 591-602. doi: 10.1016/j.apsusc.2018.01.022

28. Huang, Yukun, Lei Mei, Xianggui Chen, and Qin Wang, "Recent developments in food packaging based on nanomaterials." ***Nanomaterials,*** Vol. 8, No. 10 (2018), 830. doi: 10.3390/nano8100830

29. Siró, István, and David Plackett, "Microfibrillated cellulose and new nanocomposite materials: a review." ***Cellulose,*** Vol. 17, No.3 (2010), 459-494. doi: 10.1007/s10570-010-9405-y

30. He, Wen, Shenxue Jiang, Qisheng Zhang, and Mingzhu Pan, "Isolation and characterization of cellulose nanofibers from Bambusa rigida." ***BioResources,*** Vol. 8, No. 4 (2013), 5678-5689. doi: 10.15376/biores.8.4.5678-5689

31. Yudhanto, F, Jamasri and Rochardjo, H.S.B., "Physical and Thermal Properties of Cellulose Nanofibers (CNF) Extracted from Agave Cantala Fibers Using Chemical-Ultrasonic Treatment." ***International Review of Mechanical Engineering,***

Vol. 12, No. 7, (2018), 597-603. doi: 10.15866/ireme.v12i7.14931

32. Yudhanto, F., Jamasri and Rochardjo, H.S.B., "Physical and Mechanical Characterization of Polyvinyl Alcohol Nanocomposite Made from Cellulose Nanofibers." In ***Materials Science Forum***, Trans Tech Publications Ltd. Vol. 988, (2020), 65-72. doi: 10.4028/www.scientific.net/MSF.988.65
33. Segal, L., Creely, J. J., Martin Jr., A.E., Conrad, C.M., "Anempirical method for estimating the degree of crystallinity ofnative cellulose using the X-ray diffractometer." ***Textile Research Journal,*** (1959), 786-794. doi: 10.1177/004051755902901003
34. Schick, C., "Differential scanning calorimetry (DSC) of semicrystalline polymers." ***Analytical and Bioanalytical Chemistry,*** Vol. 395, No. 6 (2009), 1589. doi: 10.1007/s00216-009-3169-y
35. Patel, Arunendra Kumar, R. Bajpai, and J. M. Keller, "On the crystallinity of PVA/palm leaf biocomposite using DSC and XRD techniques." ***Microsystem Technologies,*** Vol. 20, No.1 (2014), 41-49. doi: 10.1007/s00542-013-1882-0
36. Kong, Y., and J. N. Hay, "The measurement of the crystallinity of polymers by DSC." ***Polymer,*** Vol. 43, No.14 (2002), 3873-3878. doi: 10.1016/S0032-3861(02)00235-5
37. Rosa, M. F., Medeiros, E. S., Malmonge, J. A., Gregorski, K. S., Wood, D. F., Mattoso, L. H. C., ... & Imam, S. H, "Cellulose nanowhiskers from coconut husk fibers: Effect of preparation conditions on their thermal and morphological behavior." ***Carbohydrate Polymers,*** Vol. 81, No. 1 (2010), 83-92. doi: 10.1016/j.carbpol.2010.01.059
38. Shahbazi, P., Behzad, T., & Heidarian, P, "Isolation of cellulose nanofibers from poplar wood and wheat straw: optimization of bleaching step parameters in a chemo-mechanical process by experimental design", ***Wood Science and Technology,*** Vol. 51, No. 5, (2017), 1173-1187 doi: 10.1007/s00226-017-0929-2
39. Neto, W. P. F., Mariano, M., Da Silva, I. S. V., Silvério, H. A., Putaux, J. L., Otaguro, H., ... & Dufresne, "A. Mechanical properties of natural rubber nanocomposites reinforced with high aspect ratio cellulose nanocrystals isolated from soy hulls." ***Carbohydrate Polymers,*** Vol. 153, (2016), 143-152. doi: 10.1016/j.carbpol.2016.07.073
40. Hassan, M., Berglund, L., Hassan, E., Abou-Zeid, R., & Oksman, K., "Effect of xylanase pretreatment of rice straw unbleached soda and neutral sulfite pulps on isolation of nanofibers and their properties." ***Cellulose,*** Vol. 25, No.5 (2018), 2939-2953. doi: 10.1007/s10570-018-1779-2
41. de Morais Teixeira, E., Bondancia, T. J., Teodoro, K. B. R., Corrêa, A. C., Marconcini, J. M., & Mattoso, L. H. C., "Sugarcane bagasse whiskers: extraction and characterizations." ***Industrial Crops and Products,*** Vol. 33, No.1 (2011), 63-66. doi: 10.1016/j.indcrop.2010.08.009
42. Siqueira, G., Tapin-Lingua, S., Bras, J., da Silva Perez, D., & Dufresne, A, "Morphological investigation of nanoparticles obtained from combined mechanical shearing, and enzymatic and acid hydrolysis of sisal fibers." ***Cellulose,*** Vol. 17, No. 6 (2010), 1147-1158. doi: 10.1007/s10570-010-9449-z
43. Poletto, Matheus, Vinícios Pistor, and Ademir J. Zattera, "Structural characteristics and thermal properties of native cellulose." ***InTech,*** Vol. 10, (2013), 2705. doi: 10.5772/50452. 1
44. Khalil, HPS Abdul, H. Ismail, H. D. Rozman, and M. N. Ahmad, "The effect of acetylation on interfacial shear strength between plant fibres and various matrices." ***European Polymer Journal,*** Vol. 37, No. 5 (2001), 1037-1045. doi: 10.1016/S0014-3057(00)00199-3
45. Rezaei, S., Najafpour, G.D., Mohammadi, M. Moghadamnia, A.A. Kazemi., "Formic acid and microwave assisted extraction of curcumin from turmeric (*Curcuma longa* L.)." ***International Journal of Engineering. Transactions B: Applications,*** Vol. 29, No. 2 (2016), 145-151. doi: 10.5829/idosi.ije.2016.29.02b.02
46. Lim, Lim, Noor Rosli, Ishak Ahmad, Azwan Mat Lazim, and Mohd Mohd Amin, "Synthesis and swelling behavior of pH-sensitive semi-IPN superabsorbent hydrogels based on poly (acrylic acid) reinforced with cellulose nanocrystals." ***Nanomaterials,*** Vol. 7, No. 11 (2017), 399. doi: 10.3390/nano7110399
47. Sadeghifar, Hasan, and Arthur Ragauskas. "Lignin as a UV Light Blocker—A Review." ***Polymers,*** Vol. 12, No. 5 (2020): 1134. doi: 10.3390/polym12051134
48. Agrawal, lS L., and Arvind Awadhia, "DSC and conductivity studies on PVA based proton conducting gel electrolytes." ***Bulletin of Materials Science,*** Vol. 27, No. 6 (2004), 523-527. doi: 10.1007/BF02707280
49. Rynkowska, Edyta, Kateryna Fatyeyeva, Stéphane Marais, Joanna Kujawa, and Wojciech Kujawski, "Chemically and Thermally Crosslinked PVA-Based Membranes: Effect on Swelling and Transport Behavior." ***Polymers,*** Vol. 11, No.11, (2019), 1799. doi: 10.3390/polym11111799
50. Nishiyama, Yoshiharu, Paul Langan, and Henri Chanzy. "Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction." ***Journal of the American Chemical Society,*** Vol. 124, No. 31 (2002), 9074. doi: 10.1021/ja0257319

Persian Abstract

چکیده

در این مقاله یک مطالعه تجربی از افزودن نانوالیاف سلولزی (CNF) استخراج شده توسط فرآیند فراصوت شیمیایی از گیاهان برگ آگاو کانتالا در ماتریس پلی وینیل الکل (PVA)ارائه شده است. با ترکیب این مواد ، فیلم نانوکامپوزیتی با ضخامت 30 میکرومتر تولید می شود. ویژگی نانوکامپوزیت با افزودن CNF (0، 2 ، 5 ، 8 و 10 درصد وزنی) در تعلیق (3PVA درصد وزنی) بررسی شد. فیلم های نانوکامپوزیتی PVA / CNF با روش محلول ریخته گری تهیه شد. رشته های الیاف به CNF با استفاده از میکروسکوپ الکترونی روبشی و میکروسکوپ الکترونی عبوری مورد تجزیه و تحلیل قرار گرفت. پیوند شیمیایی مولکولی و تجزیه و تحلیل ساختاری گروه عملکردی فیلم نانوکامپوزیت با استفاده از پراش مادون قرمز و تبدیل اشعه ایکس تبدیل فوریه مورد آزمایش قرار گرفت. فیلم نانوکامپوزیت PVA / CNF دارای مزایای قابل توجهی در مانع ماورا بنفش ، پایداری حرارتی آزمایش شده توسط کالریمتری اسکن دیفرانسیل و تجزیه و تحلیل گرما و اندازه گیری مقاومت است. به طور کلی ، افزودن بهینه CNF در ماتریس 8 درصد وزنی است ، در نتیجه بالاترین شاخص تبلور (37/5 درصد) ، مقاومت کششی و کشیدگی در شکست به ترتیب 79 و 138 درصد افزایش داشت. دارای اشعه ماورا بنفش جذب کننده خوب (82/4 درصد) و پایداری حرارتی بالا (365 درجه سانتیگراد) است.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 999-1007

International Journal of Engineering

Journal Homepage: www.ije.ir

Experimental Damage Evaluation of Honeycomb Sandwich with Composite Face Sheets under Impact Load

D. Muniraj*, V. M. Sreehari

School of Mechanical Engineering, SASTRA Deemed to be University, Tamilnadu, India

PAPER INFO

Paper history:
Received 26 August 2020
Received in revised form 24 January 2021
Accepted 03 February 2021

Keywords:
Damage
Honeycomb Sandwich
Composite
Impact
Compression after Impact

ABSTRACT

Aerospace structures are highly vulnerable to impact loads whose damage tolerance, and its resistance vary over the range of impact velocity. Honeycomb sandwich structures are used in aerospace industries where mass efficient and impact resistant structures are needed. However, the structural integrity of these structures is reduced by impact load due to tool drop, runway debris, hailstones and improper handling of the structure. A thorough investigation of the damage behaviour of honeycomb sandwich under low-velocity impact and the post-impact residual strength determination is required to design a crashworthy lightweight structure. This paper presents the experimental evaluation of specific energy absorption using Charpy impact, residual compressive strength by compression after impact and damage evaluation of honeycomb sandwich structures having composite face sheets. Parametric studies on composites and honeycombs are carried out by varying the cell size, cell thickness, core height, impact velocity, thickness and orientation of lamina. Densely packed thick honeycombs provide higher fracture energy. Under transverse compressive loading, the honeycomb core undergoes cell wall buckling, crushing and densification. Load-displacement history under in-plane compression and compression after impact for different impact energies is observed. The present study contributes for the understanding how various parameters affect the characteristics of face sheet indentation and plastic buckling of honeycomb sandwich structures with composite face sheets, thereby providing useful guidelines for its potential applications in impact engineering.

doi*: 10.5829/ije.2021.34.04a.26*

1. INTRODUCTION

Honeycomb sandwich structures with composite face sheets are increasingly used in aerospace, automobile and marine industries due to their combined excellent mechanical properties and lightweight. For example, flight control surfaces, engine cowling, helicopter rotor blades, ship hulls, automotive rims, chassis components and spoilers are made up of honeycomb sandwich structures, which hugely reduce the vehicle weight and fuel consumption. These structures are subjected to low-velocity impact due to tool drop during maintenance, runway debris and also high-velocity impact due to hailstones, bird hit and micrometeoroids. This event leads to indentation, perforation, skin fracture, core detachment and complete penetration based on the impact energy. The compressive strength of the structure drastically reduces after an impact event. To avoid accidents and to endure impact loads, many lightweight materials that absorb energy and having very high strength to weight ratio are used as shields for aerospace structures. In recent years, structural crashworthiness has become a significant area of study for the benefits of public safety and social economy. Hence, it is essential to design a structure which is crashworthy and lightweight considering the safety of the crew and the cost of the mission. The solution for attaining these two competing parameters is by using the honeycomb sandwich composite structure.

Sandwiched composite belongs to a special family of composite materials that consists of two thin face sheets bonded to the top and bottom of the lightweight core, which performs like an I beam web to sustain the shear load and transfer load to the face sheets placed away from

*Corresponding Author Institutional Email: muniraj@sastra.ac.in (D. Muniraj)

Please cite this article as: D. Muniraj, V. M. Sreehari, Experimental Damage Evaluation of Honeycomb Sandwich with Composite Face Sheets under Impact Load, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 999-1007

the neutral axis. Composites made of glass, carbon, aramid fibers reinforcement with epoxy resin, polypropylene, and polyamide matrix are used as face sheets in sandwiched structure due to its high strength to weight ratio. Composites have high load-bearing and impact resistance in the fiber direction. Aluminium, steel, nomex honeycomb and few low-density foams are widely used as the core material in the sandwiched structure. Green composites are one of the important class of composite materials made from biodegradable polymers and natural fibers as reinforcement, which are used in the bicycle frame, fork, dashboard, false ceiling, safety helmet, wind turbine and cross arms in transmission tower [1-3]. The hybridisation of glass fibers with other high strength eco-friendly fibers like basalt fiber, treated sugar palm offers better mechanical and thermal properties [4, 5]. Aisyah et al. [6] studied the effect of hybridisation on thermal properties of carbon fiber with woven kenaf reinforced epoxy composite. Maraki et al. [7] performed the design of experiment on AZ31 magnesium alloy under Charpy impact test to investigate the failure energy. Steel-concrete composite beams, frames with hybrid connectors are widely used in civil structures like multi-storeyed buildings and bridges [8, 9]. Lightweight foamed concrete is a cellular concrete manufactured in cement-based slurry by combining foam with the desired density, which is hugely used for thermal and acoustic insulation [10]. Jinan et al. [11] characterised the axial compressive failure mode of composite concrete-steel plate shear wall by bucking, cracking and crushing. Saeed et al. [12] interpreted the dependence of structural performance and its matching inter-story drift in RC moment frames. Aisyah et al. [13] elaborated the fundamentals, prospects and present view of the creep test rig for a composite cantilever beam.

Sohel et al. [14] reported the damage characteristics and performance of sandwich beams with different spacing of shear connector under impact load. Cote et al. [15] reported the manufacturing route of metallic honeycombs and it's out of plane compressive behaviour. Sibeaud et al. [16] experimentally tested the honeycomb using a two-stage light gas gun at hypervelocity normal and oblique impact ranging from 2 to 10 km/s. Jankowiak et al. [17] studied the changes in failure mode for different projectile shape due to stress triaxiality state. Xie et al. [18] implemented the interlaminar damage models and crushable foam model to simulate the low-velocity impact event of a foam core sandwich panel. The contact duration required for the entire structure to respond is more in low-velocity impact (LVI); as a result, more amount of energy is absorbed elastically. Experimental and numerical simulations have been performed to understand the impact behaviour of S-glass polyester composite laminate plate under low energy impact. Hashin's failure criteria have been used to study the inter-laminar stresses and the delamination of composites by Zouggar et al. [19]. A numerical and experimental set of low energy impact tests was carried out on composite plates in a bending configuration. Compression after impact (CAI) has been performed experimentally by Zhang et al. [20] on woven carbon fibre-reinforced composite to examine the residual compressive strength at different temperature levels. Microscopic observations have been performed by Elias et al. [21] to study the damage mechanism due to LVI in 3D woven composite employing X-ray tomography. Schroder et al. [22] discussed the usage of crushable shield beneath the lander platform which is made up of sandwich structures with aluminium core numerically and validated experimentally. Mars et al. [23] investigated the response of LVI on glass fiber reinforced polyamide using ABAQUS and correlated with the experimental results. Recently, Palomba et al. [24], used multi-walled honeycomb sandwich structures to increase the energy absorption capabilities. High-velocity impact (HVI) was carried out on aluminium honeycomb sandwich panels using a gas gun and measured the residual velocity, energy absorption and its internal damage using X-ray tomography and 3D scanning by Sun et al. [25]. Honeycomb sandwich structure under HVI at elevated temperature was experimentally and numerically studied by Xie et al. [26]. Mertani et al. [27] observed that honeycombs have excellent energy absorption properties when an impact load is applied in the out of plane direction which progressively causes cell wall buckling, core crushing and densification of the core. Babaei [28] evaluated the experimental responses of the clamped mild steel, copper and aluminium circular plates under blast loading.

Extensive research was conducted in the field of impact engineering, but limited literature was reported for the study of post-impact behaviour and its residual strength. However, a thorough understanding of post-impact damage and CAI response is still required to improve the crashworthiness of the structure under various impact energies. This paper presents an experimental investigation of honeycomb sandwich with composite face sheet under impact load and the related fracture energy, damage behaviour and residual compressive strength. A parametric study was performed to study the toughness through fracture energy using Charpy impact. The reduction of compressive strength of impacted specimen at different impact energies through CAI was compared with the non-impacted specimen.

2. EXPERİMENTAL INVESTIGATION

2. 1. Materials Honeycomb sandwich structures are made with aluminium honeycomb AA 3003 and glass fiber composite face sheets. Honeycombs having various core height, cell size, cell thickness and E glass fibrous composite made of random and preferred orientations are used to study the impact resistance and residual compressive strength. Standard epoxy resin LY 556 and hardener HY 951 was used for sandwich construction. Aluminium honeycombs and composite constituents were procured from Eco Earth Solutions Pvt Ltd, Mumbai and New Era Composites, Chennai, respectively.

2. 2. Manufacturing Vacuum-assisted resin transfer mould VARTM technique was used to fabricate the composite plate. E glass bidirectional fibre having 200 gsm was cut and layed up one over another on the 30 cm x 30 cm mould (Figure 1) as per the requirement. LY 556 standard epoxy resin and HY 951 hardener in the ratio of 10:1 was mixed properly without bubble formation. The fiber-resin volume fraction of 60:40 is maintained for manufacturing. Polyvinyl coating is applied on the surface of the mould and allowed to cure for non-sticky purpose. The fibres are placed on the mould over which peel ply resin infusion mesh and breather is packed as vacuum bag. Vacuum pressure of 0.5 bar is maintained inside the bag for 15 minutes and checked for leakage. Then the resin hardener mixture is allowed to infuse and spread all over the fiber. Then the complete set up is allowed to cure for 12 hours in the atmospheric temperature. AA 3003 Aluminium honeycombs are used for the sandwich construction. Nine sandwich panels having three distinct stacking sequences, skin thickness and ply orientations are shown in Figure 2. For sandwich construction, the face sheets are bonded to the honeycomb using the resin hardener mixture.

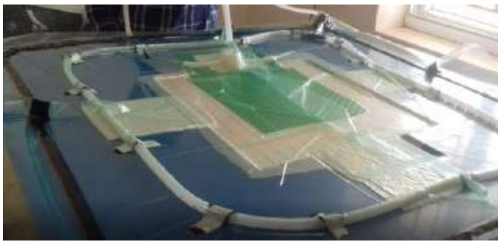


Figure 1. Vacuum-assisted resin transfer mould

Figure 2. Constructed sandwich panel

2. 3. Charpy Impact Test Charpy impact is a pendulum type of low-velocity impact test in which the mass is raised to a height and released to swing about the pivot. The pendulum strikes the specimen, eventually fractures at a higher strain rate, during which the total energy absorption E_t can be evaluated. The fracture energy E* is used to study the fracture toughness properties of the sandwich structure.

2. 4. Freefall Impact Honeycombs have excellent energy absorption capabilities under impact load. Internally gained energy is dissipated through elastic deformation, permanent plastic deformation, heat and sound. A spherical cast iron ball is dropped from a height (h) in a hollow tube which produces low-velocity impact damage on the localized area of the test specimen. The damage mainly depends upon the mass and velocity of the indenter. A mass of 3.7 kg (S1) and 2.8 kg (S2) ball is dropped from a height of 1.3 m, creates an impact on 100 mm x 100 mm honeycomb sandwich structure.

2. 5. Compression After Impact Composites are highly prone to impact damage; even a blow with impact energy of 1 J creates irreversible damage. This operation damage may not be visible for the naked eye, but this could lead to complete failure during the in-service when other loads are acting. Especially, when the compressive load acts on a structure after an impact event, the compressive strength decreases. The specimen undergoes compression test after the impact event to evaluate the residual compressive strength. The impacted specimen using free-fall impact is cut into 100 mm x 50 mm with

TABLE 1. Fracture energy for different core parameters

Sample ID	Cell Size (mm)	Core Height (mm)	Cell Thickness (µm)	Composite Thickness (mm)	C.S area (mm^2)	E_t (J)	E^* (KJ/m^2)
H_1	3	8	50	1.5	110	7	63.63
H_2	6	10	50	1.5	130	5	38.46
H_3	9	21.5	50	1.5	230	7	30.43
T_1	6	16	80	1.5	190	5	26.31
T_2	6	16	70	1.5	190	4	21.05

impact location as center and in-plane compressive load is applied on the honeycomb sandwich structure.

2. 6. Out of Plane Compression Out of plane compressive load is applied to the sandwich structure to examine the failure load of the honeycomb. It is assumed that the composite face sheet bonded to the honeycomb has a negligible effect on failure load since Young's modulus of the honeycomb core is less than the face sheet. The compressive load is applied on the top face sheet of the 50 mm x 50 mm sandwich composite test specimen.

3. RESULTS AND DISCUSSION

3. 1. Effect of Geometrical Parameters of Honeycomb Core The standard Charpy test specimen of 80 mm length having three different cell size and core height H1, H2, H3 and two different core thicknesses T1, T2, before and after the damage is shown in Figure 3 and less deterioration is observed in H1. The toughness of a material is based on the energy absorption and plastic deformation, typically area under the stress-strain curve before failure. Toughness property of a sandwich structure is studied in terms of fracture energy E*. It is evident from Table 1 that, the increase in honeycomb cell size and core height results in the reduction of fracture energy and the increase in cell thickness results in the increase of fracture energy. The core is densely packed in sample H1 and hence relatively an increased value of fracture energy is observed. As the cell thickness increases, the cell wall buckling characteristics improve, due to which the fracture energy increased.

3. 2. Effect of Skin Thickness and Ply Orientation of the Composite Plate The drop tests for low-velocity impact are performed in a guided mass falling through a tube having 1 m height and eventually impacts the 100 mm x 50 mm sandwich panel. The impactor has a spherical shape with a 3.90 kg of weight. The free fall height, as well as the weight of the impactor, is modifiable to allow testing in a large energy range. Inplane compressive load is applied at the two edges of the impacted and non-impacted sandwich structure. The test fixture (Figure 4) is made as per the standard to conduct the compression test after impact.

The sandwich structure is supported at the two edges with the test fixture to avoid slipping and global buckling. The compression test is carried out in a universal testing machine (Figure 5) at constant strain rate. The composite plate failed in the middle and the crushing of the honeycomb core is shown in Figure 6. The compressive strength variation for two different skin thickness of 1.5 mm and 2.5 mm using 0/90 combinations of ply orientation for the sandwich structure is shown in Figure 7 a1, a2, b1 and b2. As the thickness of the composite skin is increased, the compressive strength increases for both impacted and non-impacted specimen. Upon comparing Figure 7 a1 and a2 the compressive strength decreases due to the damage in the impacted specimen. For damaged specimen, the displacement corresponding to the peak compressive strength is higher than the undamaged specimen. A positive drift is observed

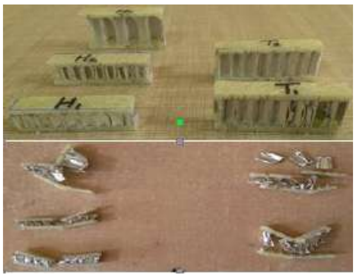


Figure 3. Sandwich structure before and after Charpy impact

Figure 4. Test fixture for CAI

Figure 5. Specimen under compression in UTM with the test fixture

Figure 6. Damaged honeycomb sandwich after CAI

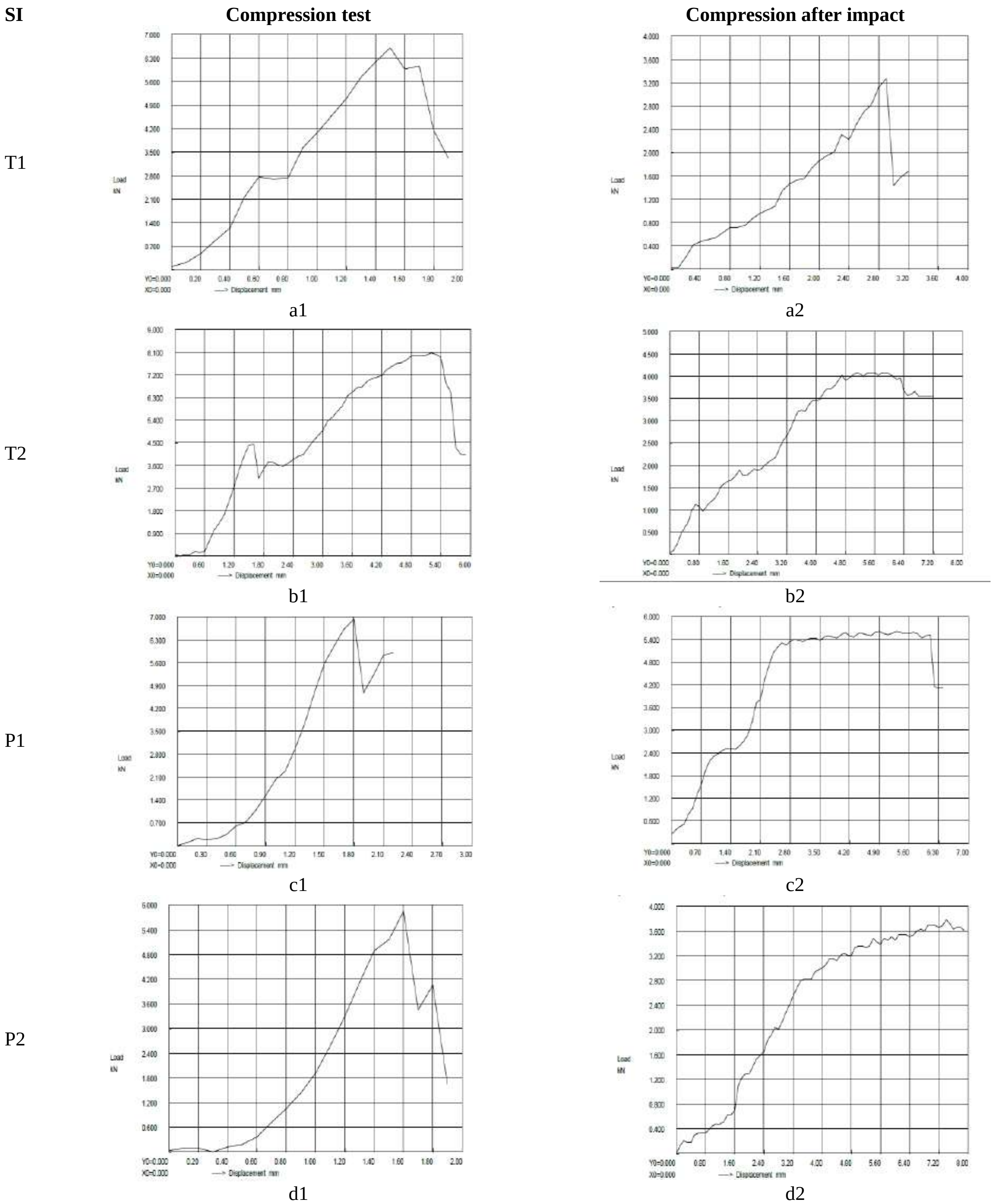


Figure 7. Load (in kN) versus displacement (in mm) curves showing compressive strength variation in impacted and non-impacted specimen

between Figure 7 a1 and a2. Figure 7 c1, c2, d1and d2 shows the compressive strength variation for two different ply orientation at 0/90/30/45 combinations on the 3 mm thick honeycomb sandwich structure. The ply orientation P1 (0/90/45/-45)4 having 45-degree plies provides better compressive strength than 30-degree plies in P2 (0/90/30/-30)4. A sudden drop of load after the peak is observed in the undamaged specimen but occurs at a higher value of displacement owing to the enduring capability, whereas the load gradually varies in the damaged specimen at constant strain rate.

To study the transverse load-bearing capacity, out of plane compressive load is applied on the surface of the composite skin. Since the modulus of elasticity of the honeycomb core is less than the composite, initially the honeycomb core cell wall buckles and eventually fails by core crushing. Beyond straining this level, the core bulges in the sideward direction, so the ram of the UTM takes the load and hence the exponential variation is observed in the curve (Figure 8). The compressive strength reduction due to impact event has a huge dependency on ply orientation (Table 2).

3. 3. Damage under Compression and CAI at Various Impact Energy

As the spherical ball impacts the top face sheet, barely visible damage occurs and its fracture pattern is highlighted on the specimen. The fracture occurs at the edges of the dent diameter of the ball in the sample S1 for impact energy of 47 J and slight fracture occurs at the center in sample S2 for impact energy of 36 J as shown in Figures 10 a1 and b1, respectively. The cell wall buckling and the permanent deformation on the top of the honeycomb are shown in Figure 9. The impacted specimens (S1 and S2) and non-impacted specimen (S3) is compressed inplane and the corresponding damage is shown in Figures 10 a2, b2 and c2, respectively. The non-impacted specimen has not failed completely, whereas the damage occurs at the clamped edges and propagates progressively.

The damage due to out of plane compression in specimen C1 is shown in Figure 10 d2, in which the core deteriorated completely due to its lower modulus of elasticity when compared to the composite. The compressive force increases gradually reach its maximum and then reduces suddenly once the face sheet fails due to buckling, in the impacted specimens as shown in Figure 11 a and b. For non-impacted specimen, the curve reaches its maximum and remains the same with minor fluctuation even for the larger deformation (Figure 11 c). The compressive strength of the impacted specimen decreases when compared to the non-impacted specimen. Even though the impact energy of the specimen S1 is higher, the percentage reduction in compressive strength is lower than the specimen S2 (Table 3). This is due to the fact, that the contact area for S1 is larger due to the large diameter spherical ball. Under the in-plane compression, displacement increases till cell wall buckling and remains constant during core crushing. Finally, it reaches the densification phase where the load rapidly increases for smaller deformation.

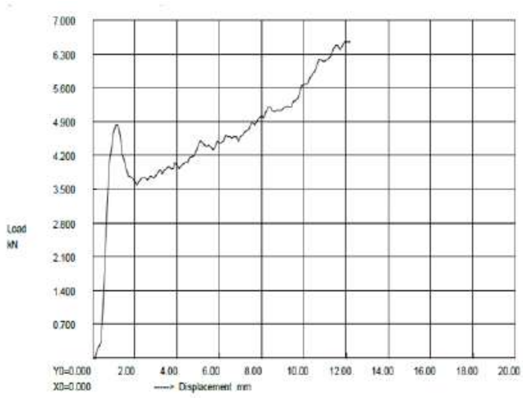


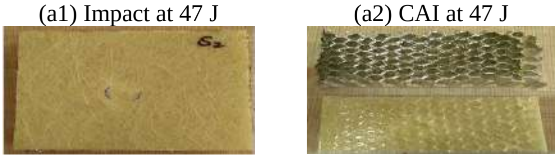
Figure 8. Out of plane compression test (Load (in kN) versus displacement (in mm) curve)

TABLE 2. Compressive strength reduction due to impact

Specimen	Thickness (mm)	Orientation (Degree)	Compression test (KN)	Compression after impact (KN)	Percentage reduction in compressive strength
T1	1.5	$(0/90)_4$	6.7	3.3	50.74
T2	2.5	$(0/90)_6$	8.1	4.1	49.38
P1	3	$(0/90/45/-45)_4$	6.9	5.5	20.28
P2	3	$(0/90/30/-30)_4$	5.8	3.7	36.20

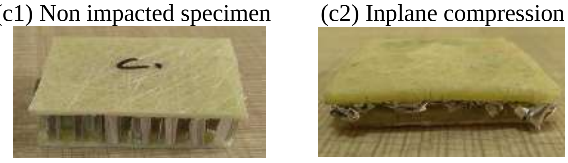
Figure 9. Plastic deformation of honeycomb

(a1) Impact at 47 J (a2) CAI at 47 J



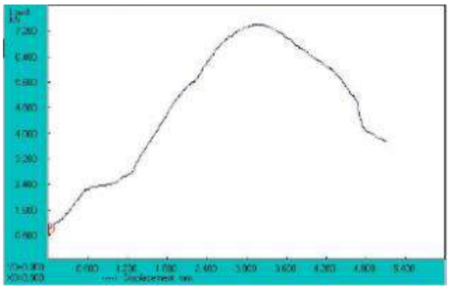
(b1) Impact at 36 J (b2) CAI at 36 J

(c1) Non impacted specimen (c2) Inplane compression

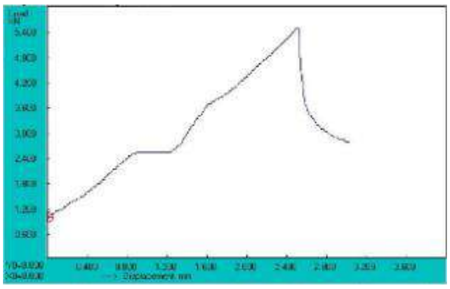


(d1) Without impact (d2) Out of plane compression

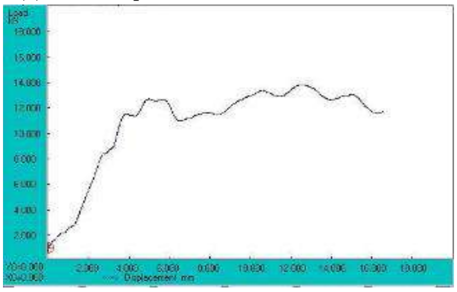
Figure 10. Damaged pictures of honeycomb sandwich panels under various loading condition



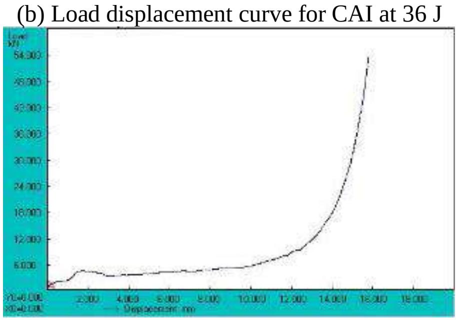
(a) Load displacement curve for CAI at 47 J



(b) Load displacement curve for CAI at 36 J



(c) Load displacement curve for non impacted specimen



(d) Out of plane compression

Figure 11. Typical load (in kN)-displacement(in mm) curve under various loading

TABLE 3. Compressive strength reduction for different impact energies

Specimen ID	Maximum force Fmax (KN)	Displacement at Fmax (mm)	Compressive strength (MPa)	Percentage reduction of compressive strength
S1	7.4	3.09	6	50
S2	5.5	2.49	5	58.3
S3	13.8	12.38	12	-
C1	12	14	4.8	-

4. CONCLUSIONS

The ultimate goal of achieving the high strength/stiffness to weight ratio and crashworthy structure for the aerospace applications is attained through honeycomb sandwich composite structures. The significant conclusions made from the present study are mentioned below.

- Densely packed cells with higher cell thickness possess higher fracture energy E*. The contact area is also one of the main significant parameters that decide the damage.
- Due to an impact event, the compressive strength decreases to an enormous degree. Hence significant attention has to be paid for CAI, which may cause

failure during an in-service operation before it reaches the ultimate load.

- Out of plane compression test reveals that the honeycomb undergoes progressive damage and hence honeycomb core with composite face sheets can be used as an energy-absorbing sandwich structure to withstand impact load.
- Thus the present study contributes for the understanding how various parameters affect the characteristics of face sheet indentation and plastic buckling of honeycomb sandwich structures with composite face sheets, which are highly significant in several real-world applications. Present study has been carried out at room temperature, but the study may be extended to analyze the effect at higher temperature, which will be useful for aerospace applications.

5. ACKNOWLEDGEMENT

This research work was supported by the Science Engineering Research Board (SERB), Department of Science & Technology (DST), Government of India (Grant No. ECR/2018/001065) and Prof. T R Rajagopalan fund, SASTRA Deemed to be University. Authors express sincere thanks to the support from SERB-DST and SASTRA Deemed to be University.

6. REFERENCES

1. Mekonnen, B.Y. and Mamo, Y. J., "Tensile and Flexural Analysis of a Hybrid Bamboo/Jute Fiber-reinforced Composite with Polyester Matrix as a Sustainable Green Material for Wind Turbine Blades", ***International Journal of Engineering, Transactions B: Applications***, Vol. 33, No. 2, (2020), 314-319. https://doi.org/10.5829/IJE.2020.33.02B.16
2. Asyraf, M. R. M., Ishak, M. R., Sapuan, S. M., Yidris, N., Ilyas, R.A., Rafidah, M. and Razman, M. R., "Potential Application of Green Composites for Cross Arm Component in Transmission Tower: A Brief Review", ***International Journal of Polymer Science***, (2020), 8878300, https://doi.org/10.1155/2020/8878300.
3. Asyraf, M. R. M., Ishak, M. R., Sapuan, S. M., Yidris, N., Ilyas, R.A., Rafidah, M. and Razman, M. R., "Evaluation of Design and Simulation of Creep Test Rig for Full-Scale Crossarm Structure", ***International Journal of Polymer Science***, (2020), 6980918, https://doi.org/10.1155/2020/6980918.
4. Nurazzi, N. M., Khalina, A., Sapuan, S. M., Ilyas, R. A., Rafiqah, S. A. and Hanafee, Z. M., "Thermal properties of treated sugar palm yarn/glass fiber reinforced unsaturated polyester hybrid composites", ***Journal of Materials Research and Technology***, Vol. 9, No. 2, (2020), 1606-1618. https://doi.org/10.1016/j.jmrt.2019.11.086.
5. Sapuan, S. M., Aulia, H. S., Ilyas, R. A., Atiqah, A., Dele-Afolabi, T. T., Nurazzi, M. N. and Atikah, M. S. N., "Mechanical properties of longitudinal basalt/woven-glass-fiber-reinforced unsaturated polyester-resin hybrid composites", ***Polymers***, Vol. 12, No. 10, (2020), 1-14. https://doi.org/10.3390/polym12102211.
6. Aisyah, H. A., Paridah, M. T., Sapuan, S. M., Khalina, A., Berkalp, O. B., Lee, S. H. and Ilyas, R. A., "Thermal Properties of Woven Kenaf/Carbon Fibre-Reinforced Epoxy Hybrid Composite Panels", ***International Journal of Polymer Science,*** (2019). https://doi.org/10.1155/2019/5258621.
7. Maraki, M. R., Tagimalek, H., Azargoman, M., Khatami, H. and Mahmoodi, M., "Experimental investigation and statistical modeling of the effective parameters in Charpy impact test on AZ31 magnesium alloy with v-shape groove using Taguchi method", ***International Journal of Engineering Transactions C: Aspects***, Vol. 33, No. 12, (2020), 2521-2529. https://doi.org/10.5829/ije.2020.33.12c.13.
8. Lu, S., Zhao, W., Han, P. and Hang, Z., "Mechanical Behavior of Hybrid Connectors for Rapid-Assembling Steel-Concrete Composite Beams", ***Civil Engineering Journal***, Vol. 5, No. 10, (2019), 2081-2092. https://doi.org/10.28991/cej-2019-03091395.
9. Kuddus, M. A. and Dey, P. P., "Cost Analysis of RCC, Steel and Composite Multi-Storied Car Parking Subjected to High Wind Exposure in Bangladesh", ***Civil Engineering Journal***, Vol. 3, No. 2, (2017), 95-104. https://doi.org/10.28991/cej-2017-00000076.
10. Wani, S. B., "Influence of Bi-directional Fibreglass Grid Reinforcement on Drying Shrinkage and Mechanical Properties of Lightweight Foamed Concrete", ***International Journal of Engineering, Transactions A: Basics***, Vol. 34, No. 1, (2021), 10-18. https://doi.org/10.5829/ije.2021.34.01a.02.
11. Rastegarian, S. and Sharifi, A., "An investigation on the correlation of inter-story drift and performance objectives in conventional RC frames", ***Emerging Science Journal***, Vol. 2, No. 3, (2018), 140-147. https://doi.org/10.28991/esj-2018-01137.
12. Abbas, J. L. and Allawi, A. A., "Experimental and Numerical Investigations of Composite Concrete-Steel Plate Shear Walls Subjected to Axial Load", ***Civil Engineering Journal***, Vol. 5, No. 11, (2019), 2402-2422. https://doi.org/10.28991/cej-2019-03091420.
13. Asyraf, M. R. M., Ishak, M. R., Sapuan, S. M., Yidris, N., Shahroze, R. M., Johari, A. N. and Ilyas, R. A., "Creep test rig for cantilever beam: fundamentals, prospects and present views", ***Journal of Mechanical Engineering and Sciences***, Vol. 14, No. 2, (2020), 6869-6887. https://doi.org/10.15282/jmes.14.2.2020.26.0538.
14. Sohel, K. M. A., Richard Liew, J. Y., Alwis, W. A. M. and Paramasivam, V., "Experimental investigation of low-velocity impact characteristics of steel-concrete-steel sandwich beams", Steel and ***Composite Structures***, Vol. 3, No. 4, (2003), 289-306. https://doi.org/10.1016/S0263-8223(00)00098-2
15. Cote, F., Deshpande, V. S., Fleck, N. A. and Evans, A. G., "The out-of-plane compressive behaviour of metallic honeycombs", ***Materials Science and Engineering A***, Vol. 380, No. 1, (2004), 272-280. https://doi.org/10.1016/j.msea.2004.03.051.
16. Sibeaud, J. M., Thamié, L. and Puillet, C., "Hypervelocity impact on honeycomb target structures: Experiments and modelling", ***International Journal of Impact Engineering***, Vol. 35, No. 12, (2008), 1799-1807. https://doi.org/10.1016/j.ijimpeng.2008.07.037.
17. Jankowiak, T., Rusinek, A., Kpenyigba, K. M. and Pesci, R., "Ballistic behaviour of steel sheet subjected to impact and perforation", ***Steel and Composite Structures***, Vol. 16, No. 6, (2014), 595-609. https://doi.org/10.12989/scs.2014.16.6.595.
18. Xie, Z., Yan, Q., and Li, X., "Investigation on low-velocity impact on a foam core composite sandwich panel", ***Steel and Composite Structures,*** Vol. 17, No. 2, (2014), 159-172. https://doi.org/10.12989/scs.2014.17.2.159.
19. Zouggar, K., Boukhoulda, F. B., Haddag, B. and Nouari, M., "Numerical and experimental investigations of S-Glass/Polyester composite laminate plate under low energy impact", ***Composites***

Part B: Engineering, Vol. 89, (2016), 169-186. https://doi.org/10.1016/j.compositesb.2015.11.021.

20. Zhang, D., Fei, Q. and Zhang, P., "Drop-weight impact behaviour of honeycomb sandwich panels under a spherical impactor", ***Composite Structures***, Vol. 168, (2017), 633-645. https://doi.org/10.1016/j.compstruct.2017.02.053.
21. Elias, A., Laurin, F., Kaminski, M. and Gornet, L., "Experimental and numerical investigations of low energy/velocity impact damage generated in 3D woven composite with polymer matrix", ***Composite Structures***, Vol. 159, (2017), 228-239. https://doi.org/10.1016/j.compstruct.2016.09.077.
22. Schroder, S., Reinhardt, B., Brauner, C., Gebauer, I. and Buchwald, R., "Development of a Marslander with crushable shock absorber by virtual and experimental testing"' ***Acta Astronautica***, Vol. 134, (2017), 65-74. https://doi.org/10.1016/j.actaastro.2017.01.023.
23. Mars, J., Chebbi, E., Wali, M. and Dammak, F., "Numerical and experimental investigations of low-velocity impact on glass fiber-reinforced polyamide" ***Composites Part B: Engineering,*** Vol. 146, (2018), 116-123. https://doi.org/10.1016/j.compositesb.2018.04.012.
24. Palomba, G., Epasto, G., Crupi, V. and Guglielmino, E., "Single and double-layer honeycomb sandwich panels under impact loading", ***International Journal of Impact Engineering***, Vol. 121, (2018), 77-90. https://doi.org/10.1016/j.ijimpeng.2018.07.013.
25. Sun, G., Chen, D., Wang, H., Hazell, P. J. and Li, Q., "High-velocity impact behaviour of aluminium honeycomb sandwich panels with different structural configurations", ***International Journal of Impact Engineering***, Vol. 122, (2018), 119-136. https://doi.org/10.1016/j.ijimpeng.2018.08.007.
26. Xie, W. H., Meng, S. H., Ding, L., Jin, H., Du, S. Y., Han, G. K. and Chi, R. Q., "High-temperature high-velocity impact on honeycomb sandwich panels", ***Composites Part B: Engineering,*** Vol. 138, (2018), 1-11. https://doi.org/10.1016/j.compositesb.2017.06.022.
27. Mertani, B. M. B., Keskes, B. and Tarfaoui, M., "Experimental Analysis of the Crushing of Honeycomb Cores Under Compression", ***Journal of Materials Engineering and Performance***, Vol. 28, No. 3, (2019), 1628-1638. https://doi.org/10.1007/s11665-018-3852-2.
28. Babaei, H., "Prediction of deformation of circular plates subjected to impulsive loading using gmdh-type neural network", ***International Journal of Engineering-Transactions A: Basics***, Vol. 27, No. 10, (2014), 1635. https://doi.org/10.5829/idosi.ije.2014.27.10a.18.

Persian Abstract

چکیده

سازه های هوافضا در برابر بارهای ضربه ای که تحمل خسارت دارند و مقاومت آن در دامنه سرعت ضربه بسیار آسیب پذیر است. سازه های ساندویچی لانه زنبوری در صنایع هوافضا مورد استفاده قرار می گیرند که سازه های کارآمد در برابر جرم و مقاوم در برابر ضربه مورد نیاز است. با این حال ، یکپارچگی سازه ای این سازه ها به دلیل افت ابزار ، بقایای باند ، سنگ های تگرگ و کار با نامناسب سازه در اثر ضربه کاهش می یابد. بررسی دقیق رفتار آسیب ساندویچ لانه زنبوری تحت تأثیر سرعت کم و تعیین مقاومت پسماند پس از ضربه برای طراحی یک ساختار سبک خراب لازم است. در این مقاله ارزیابی تجربی جذب انرژی خاص با استفاده از ضربه شارپی ، مقاومت فشاری باقیمانده توسط فشار پس از ضربه و ارزیابی آسیب ساختارهای ساندویچ لانه زنبوری دارای ورق های ترکیبی صورت ارائه شده است. مطالعات پارامتریک در مورد کامپوزیت ها و لانه زنبوری ها با تغییر اندازه سلول ، ضخامت سلول ، ارتفاع هسته ، سرعت برخورد ، ضخامت و جهت گیری لایه ها انجام می شود. شانه های عسلی ضخیم بسته بندی شده انرژی شکستگی بالاتری را ایجاد می کنند. تحت بارگذاری فشاری عرضی ، هسته لانه زنبوری تحت کمانش ، خرد شدن و تراکم دیواره سلول قرار می گیرد. سابقه جابجایی بار تحت فشار و فشرده سازی درون صفحه ای پس از ضربه برای انرژی های مختلف ضربه مشاهده می شود. مطالعه حاضر به درک چگونگی پارامترهای مختلف بر خصوصیات تورفتگی ورق صورت و کمانش پلاستیکی ساختارهای ساندویچ لانه زنبوری با ورق های کامپوزیت صورت کمک می کند ، در نتیجه راهنماهای مفیدی برای کاربردهای بالقوه آن در مهندسی ضربه ارائه می دهد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 1008-1015

International Journal of Engineering

Journal Homepage: www.ije.ir

Experimental Investigation on the Effect a Rotational Shaft on the Thermal Behavior of a Circular Tube under Constant Heat Flux

G. Rahimi, S. Jafarmadar*, I. Mirzaie

Mechanical Engineering Department, Technical Education Faculty, Urmia University, Urmia, West Azerbaijan, Iran

PAPER INFO

Paper history:
Received 23 October 2020
Received in revised form 08 January 2021
Accepted 14 January 2021

Keywords:
Dimension Less Rotational Speed
Nusselt Number
Friction Factor
Thermal Efficiency Factor

ABSTRACT

Active and passive methods are two main mechanisms of heat transfer improvement. The active methods use external forces to improve heat transfer. This investigation evaluates the thermal and frictional behavior of a circular tube containing a rotational shaft. Constant heat flux was exerted to the circular tube. The fluid inlet and outlet temperature as well as wall temperature of tubes were measured to calculate the hat transfer coefficient. The Re (Reynolds) number was between 800-2000. Also, the dimensionless rotational speed (Rs) had the values of 1,1.5, 2, 2.5 and 3. Results revealed that the rotational shaft could increase the Nu number. Up to %18. Also, the results showed that the rotational shaft could significantly increase the pressure drop and friction factor. The maximum increment of %78 was achieved for friction factor. It was revealed that the use of rotational shaft could be more efficient at low Re numbers and low dimensionless rotational speeds. Also, it was found that by the increment of Reynolds number and being in the transient regime the efficiency of the system would improve.

doi: *10.5829/ije.2021.34.04a.27*

NOMENCLATURE

L	Length tube (m)	$\bar{h}$	Average convective heat transfer coefficient (W/m^2-K)
P_{out}	Outlet pressure (Pa)	A	Heat transfer area (m^2)
P_{in}	Inlet pressure (Pa)	Nu	Nusselt number
ΔP	Pressure drop (Pa)	Nu_s	Nusselt number of smooth tube
D_H	Hydraulic diameter (m)	f	friction factor
Re	Reynolds number	f_s	friction factor of smooth tube
q	Heat transfer rate(W)	TEF	Thermal efficiency factor
$\dot{m}$	Mass flow rate (kg/s)	V	Mean velocity (m/s)
T_{out}	Outlet temperature (K)	k_f	Conductive heat transfer coefficient (W/m^2-K)
T_{in}	Inlet temperature (K)	C.B.R	Cost per benefit ratio
T_w	Wall temperature (K)	**Greek symbols**	
T_b	Mean bulk temperature (K)	ρ	Density (kg/m^3)
		μ	Dynamic viscosity (kg/m-s)

1. INTRODUCTION

In accordance to the importance of heat transfer augmentation at the efficiency improvement of various thermal units, numerous researchers have implemented investigations on the heat transfer improvement techniques. The thermal enhancements mechanisms include two main categories namely as active and passive methods [1-3]. The passive methods are those at which no external force is used in the heat transfer augmentation techniques. However, at the active methods, the heat transfer enhancement is occurred by means of external forces exerted on the system [4, 5]. At the present investigation the influence of a rotational shaft on the

*Corresponding Author Institutional Email: *s.jafarmadar@urmia.ac.ir* (S. Jafarmadar)

Please cite this article as: G. Rahimi, S. Jafarmadar , I. Mirzaie, Experimental Investigation on the Effect a Rotational Shaft on the Thermal Behavior of a Circular Tube under Constant Heat Flux, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1008-1015

thermal and frictional behavior of a circular tube is experimentally investigated. Recently the use of a rotating elements in order to increase the heat transfer is significantly increased. This type heat transfer is commonly called as internal cooling technology and especially is used within the processes having a rotating shaft as like to internal cooling of turbines. At the following, some of the investigations which have used the rotational effect to increase the heat transfer are summarized.

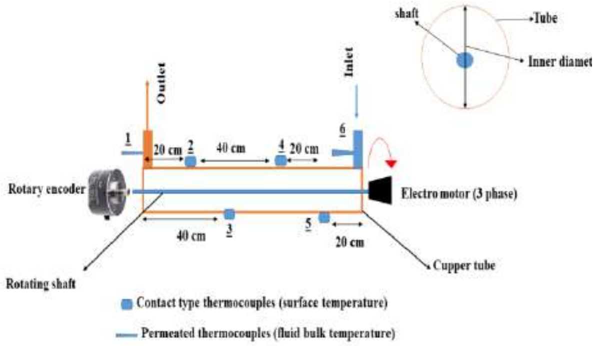
First time, White [6] reported that a rotational circular tube has better performance when compared to the stationary one. The results presented by White [6] revealed that at the higher rotational speeds, the pressure drop could be decreased up to 40%. Singaram et al. [6] and Chatterjee et al. [7] studied a half filled rotational circular tube. They proposed an empirical correlation for predicting the thickness of the film layer. Krivosheev et al. [8] probed the thermal performance of a half-filled rotational circular tube. The results revealed that as the rotational speed augments, the heat transfer rate raises. Chatterjeea et al. [9] investigated the thermal performance of a half-filled rotational circular tube. They provided an empirical correlation for the prediction of Nusselt number (Nu) inside a half-filled circular tube. Hussain and Hussein [10] implemented a numerical investigation to probe the mixed convective heat transfer within a heated square shaped enclosure at which a rotating circular shaped cylinder was running. In their study, the position of the rotating circular cylinder was the variant parameter. Park et al. [11] investigated the natural convective heat transfer within an enclosure including hot and cold rotating circular cylinder. The locating point of the hot and cold cylinder was the variant parameter which was probed by Park et al. [11]. Their results presented that as the rotating cylinder and walls of the enclosure gets closer, the heat transfer rate increases. Yoon et al. [12] numerically probed the influence of two rotating cylinder on the thermal management of a cubic. Kareem et al. [13] implemented a 3D simulation in order to study the mixed convection heat transfer within an enclosure. Selimefendigil and Öztop [14] conducted a comprehensive research to examine the effect of pulsating flow and the rotation of the cylinder on heat transfer of a heated channel. Their results revealed that the direction of the rotation and the Reynolds number (Re) could be effective on the heat transfer rate. Most of the experts studied the effect of rotational cylinder on the heat transfer, supposing that the cylinder has the circular surface. However, the triangular [15] elliptical [16] and rectangular [17] surfaces were of interest too.

Based on the above literature review and up to the author's best of knowledge, the majority of investigations have focused on the rotation of tube containing the working fluid. There are few studies that investigated the effect of a rotational element inserted in the tube. At the present paper the influence of the presence of a rotational circular shaft on the heat transfer of a circular tube under constant heat flux is studied. The Nu number, pressure drop, cost per benefit ratio (C.B.R.) and performance efficiency coefficient (PEC) factors are the parameters that are studied and investigated in this study.

2. Experimental study

2. 1. Test Rig Definition

The schematic scene of the setup is shown in Figure 1. As is presented, the test rig is consisted of a circular tube which contains a rotating shaft. The geometrical properties of the circular tube and shaft are presented in Table 1. The circular cupper made tube was put under certain value of heat flux. The constant heat flux was produced via a heater



(a) Schematic view

(b) General view

Figure 1. (a) Schematic view (b) general view of the setup

TABLE 1. Geometrical properties of the circular tube and the circular rotating shaft

Element	Inner diameter (cm)	Width (mm)	Length (m)	Thickness (mm)
Circular tube	6.6	2	1	1
Circular shaft	1.5	-	1	-

wire. With the aim of producing fixed heat flux the amount of voltage exerted to the heater wire (100 Ω resistant) was regulated via a voltage regulator (2 kW Dimmer (DS 2000VA)). A 3 cm glass wool isolator was used to reduce the heat loss. The rotary shaft was rotated by means of a 3 phase electro motor. A reductant gear box (SAHAND-W090-15:1-90B5) was also installed between the electro motor and the coupling section to reduce the rotary speed of the electro motor and to produce the wanted speeds. Additionally, the electro motor was controlled by means of an inverter (DEG Drive- DGI 900). By means of the mentioned mechanisms the accurate rotational speed was adjusted. The amount of rotational speed was measured by means of a rotary encoder (HE50B-8-1024-3-T-24) which was coupled to the end of rotational shaft. The flow outlet and inlet temperature as well as surface temperature of the circular tube was recorded to evaluate heat transfer coefficient. Six K type thermocouples (as presented in Figure 1A) with a 12 channel data logger (Lurton BTM-4208SD) were utilized for recording the fluid and surface temperature. The pressure drop (pressure drop between the points 1and 6) was evaluated by means of manometer (Lurton PM-9100). The accuracy of the data logger and manometers were ± 0.5 °C and 1 $mbar$, respectively.

2. 2. Experimental Procedure Table 2 provides different cases which are considered at the present study. As shown, four various flow rates of water flow were considered to examine the effect of Re number. At each constant water flow rate different rotational speeds were assumed too. The Rs = 0 is referred to the cases in which the circular shaft was stationary. The thermal energy produced by heater wire was kept constant during the tests. It should be noted that the initial data was noted after the system reached thermally stable condition. As is aware the repeatability and the uncertainty evaluation is urgent for any investigation. The tests were repeated for three times and then the average of these tests were used for the calculation of the investigated parameters. The uncertainty analysis was based on Moffat [18] method which was used based on reported literature [19, 20] to evaluate the uncertainty value. Table 3 presents the achieved values for uncertainty of studied parameters.

2. 3. Setup validation Furthermore to the uncertainty calculation which resulted in reasonable values, the results of the present study are compared to the existing data reported in the literature. The friction factor results were compared with the results of well-known Blasius equation. Also, the Nu number results were compared to those of the study conducted by Sun and Zeng [21]. As is presented in Figure 2, there are good agreement between the results of present study and those reported in literature. The maximum deviation was about 8% which was reasonable.

TABLE 2. various considered cases in this study

Water flow rates (l/min)	*Re*	*Rs*
2	788	0, 1,1.5, 2, 2.5, 3
3	1215	0, 1,1.5, 2, 2.5, 3
4	1519	0, 1,1.5, 2, 2.5, 3
5	1863	0, 1, 1.5, 2, 2.5, 3

TABLE 3. Uncertainty in investigated parameters

Parameter	Unit	Amount
Electrical current	mA	± 2
Electric resistence	**Ω**	± 3
Electric voltage	V	± 5
Water flow rate	lit/min	± 0.1
Water bulk temperature	℃	± 0.5
Surface temperature	℃	± 0.5
Nusselt number	%	± 9.46
Uncertainty in read values of tables ($\rho, c, \ldots$	%	$\pm 0.1 - 0.15$

3. PARAMETER DEFINITION AND CALCULATION METHOD

3. 1. Calculation Process In this part the calculation process of various parameters evaluated in this study are presented. As stated before, a digital manometer was used for measuring the pressure drop. At the following equation used for calculating pressure drop is presented.

$$\Delta P = P_{in} - P_{out} \tag{1}$$

3. 1. 1. Calculation Method of the Overall Heat Transfer Coefficient The convective heat transfer coefficient was evaluating by following method:

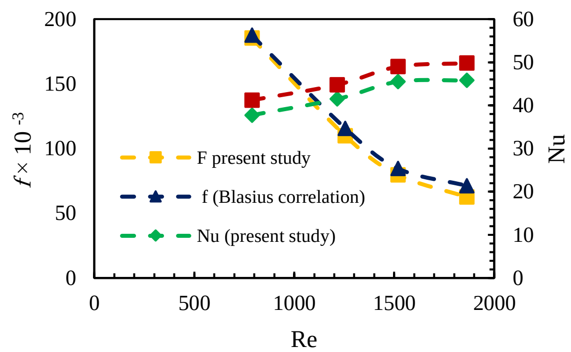


Figure 2. Comparison of the present results with existing results in the literature

As is aware, the thermal energy received by the working fluid (water flow) through the test rig is measured as follows:

$$\dot{q} = \dot{m}\, c_p (T_{out} - T_{in}) \quad (2)$$

In which the T_{in} and T_{out} are the inlet temperature and outlet temperature of the water flow, respectively.

The average heat transfer coefficient of the test rig could be calculated via the following equation.

$$\bar{h} = \frac{\dot{q}}{A\left(\frac{\sum_{j=1}^{4} T_w}{4} - T_b\right)} \quad (3)$$

where the T_b is calculated as the average of T_{out} and T_{in}.

3. 2. Parameter Definition

3. 2. 1. Friction Factor

The friction factor is the dimensionless form of the pressure drop and was calculated by the following equation.

$$f = \frac{\Delta P}{0.5\left(\frac{L}{D_h}\right)\rho V^2} \quad (4)$$

3. 2. 2. Nu Number

The Nu number denotes the potential of the heat transfer and is the dimensionless form of the heat transfer coefficient. The Nusselt number could be calculated as follows.

$$Nu = \frac{\bar{h} \times D_h}{k_f} \quad (5)$$

In which the D_h and the k_f are the hydraulic diameter and the conductive heat transfer coefficient of the fluid flow.

3. 2. 3. Cost Per Benefit Ratio Factor (C.B.R)

The Cost per benefit ratio (C.B.R) factor is defined as the increment percentage of the pressure drop per the increment percentage of the Nusselt (Nu) number. This Parameter is a criterion which could be feasibly used with engineers if they want to know that how much heat transfer improvement could provide any technique that is used.

$$C.B.R. = \frac{\%\Delta P}{\%Nu} \quad (6)$$

3. 2. 4. Performance Efficiency Coefficient (PEC) Analysis

The performance efficiency coefficient (PEC) is another criterion for evaluating the heat transfer enhancement techniques. This criterion was used widely by other experts [21, 22] to analyze the efficiency of any heat transfer augmentation method. This parameter helps engineers to choose the proper method for heat transfer augmentation. Since this parameter considers both the effect of friction factor and Nu number, it could be used for utilizing the heat transfer enhancement techniques. Performance efficiency coefficient (PEC) is defined as the ratio of Nu number of the improved tube to that of the smooth tube at the constant pumping power and is defined as below:

$$TEF = \frac{\left(\frac{Nu}{Nu_s}\right)}{\left(\frac{f}{f_s}\right)^{\frac{1}{3}}} \quad (7)$$

In which the Nu_s and f_s respectively, are the Nu number and friction factor of base tube.

3. 2. 5. Dimensionless Rotational Speed

The dimensionless rotational speed is calculated as follows:

$$R_S = \frac{\omega D_H}{2V} \quad (8)$$

4. RESULTS AND DISCUSSION

4. 1. Pressure Drop and Friction Factor Analysis

Figure 3 presents the behavior of pressure drop versus dimensionless rotational speed (Rs). As stated before, the pressure drop was recorded via a digital manometer. By looking at Figure 3, it could be concluded that, as the Rs and Re number increase, the pressure drop increases too. It could be explained that as the Rs and Re number increase the flow regime tends to transpass from laminar regime to the turbulent regime. Actually, the increment of Rs causes to generation of swirling flows which are perpendicular to the axial flow. The interactions between axial flow and the swirling flows created by rotational shaft lead in generation of more vortices and results in the pressure drop increment.

Figure 4 presents the variation of friction factor versus the Re number and the dimensionless rotational speed (Rs). As is shown, the friction factor diminishes with the increment of Re number. Although the pressure drop increases with increment of Re number (which is of first power and is at the nominator of Equation (4)) but the velocity amount increase too (which is of second power and is at the denominator). Since the effect of increment of velocity amount is dominant to the

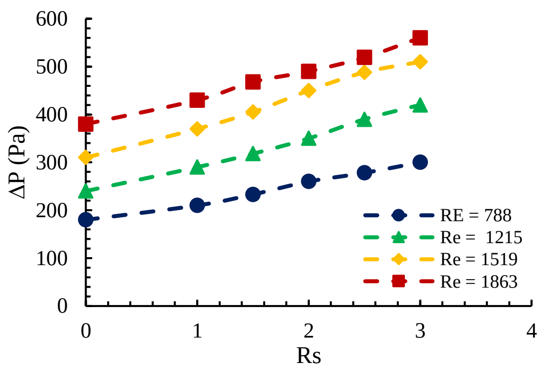


Figure 3. Pressure drop vs Rs

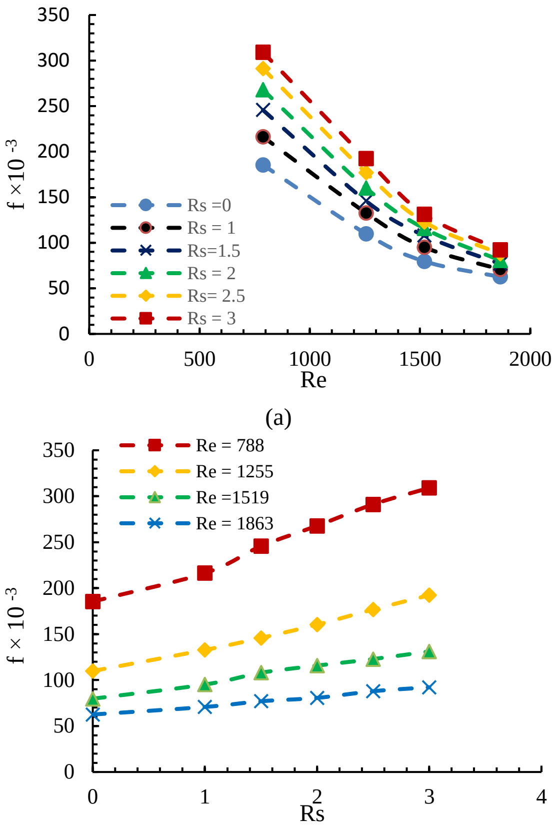


Figure 4. Friction factor coefficient vs (a) Re number (b) Rs

increment of pressure drop, as a result, the friction factor decrease. Looking at Figure 4B, it is seen that the increment of rotational speed (Rs) causes to the increment of friction factor. Indeed, the increment of Rs only increases the pressure drop and does not have effect on values of the axial velocity. So the axial velocity is independent to the amount of Rs. The maximum increment in friction factor due to increment of Rs was related to the Re = 1215 and was around 73%.

4. 2. Nusselt Number Analysis The Nusselt number (Nu) is a dimension less number and denotes the potential of heat transfer feasibility. The behavior of Nu number vs Re number and Rs is shown in Figure 5. Also, the Figure 6. presents the variation of Nu/Nu_0 vs Re number and the dimensionless rotational speed.

As is presented in Figure 5 the augmentation of both Re number and the dimensionless rotational speed cause to increment at Nu number. Indeed, that as the Re number and the Rs increases the more interactions occur between the axial flow and the swirling flows created by rotational shaft. These interactions cause to creation of eddies and vortexes. These phenomena lead in better mixing at the boundary layer. As the mixing phenomena improves the heat transfer coefficient and subsequently the Nu number increase too. Figure 6 presents the variation of Nu/Nu_0 vs Re number and Rs. As is presented the maximum increment was occurred at the lower Re numbers and the highest dimension less rotational speed. Actually, at the lower Re numbers the swirling flows could better effect on the boundary layer. As the Re number increase the axial flow is more potent than the swirling flow and could easily sweep the swirling flow with itself. Consequently, the swirling flow would have less effect on the boundary layer and less mixing phenomenon occurs at the boundary layer. Subsequently less increment at the heat transfer coefficient and the Nu number happens. Looking at Figure 6B, it could be seen that there is a deviation from trend at the highest Reynolds number (1863.98). It should be noted that the mentioned Reynolds number falls in a range that mostly in this range of Reynolds number the flow has transient regime. Within the transient regime the fluctuations in the flow intensify and the turbulence intensity of the flow is more than those related to lower Reynolds number. This increment in the turbulence intensity increase the heat transfer coefficient and consequently leads in increment in Nusselt number.

4. 3. Performance Efficiency Coefficient (PEC) Performance efficiency coefficient (PEC) is one of the

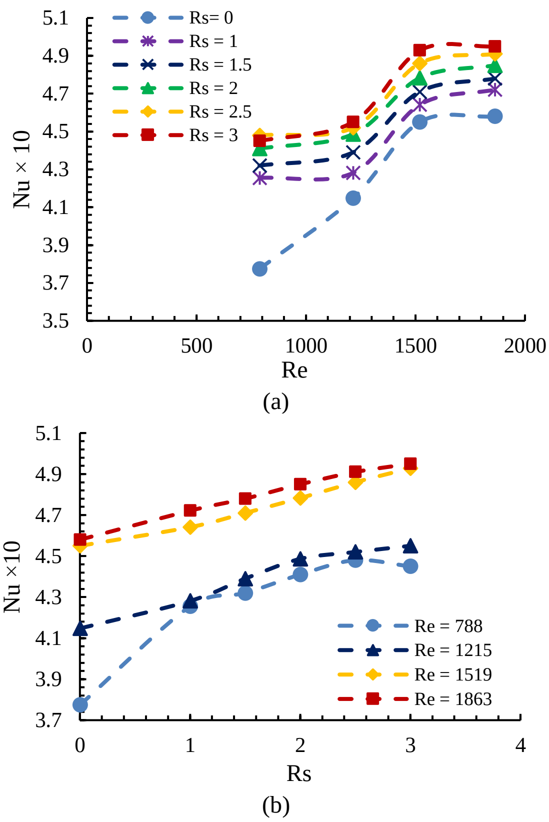
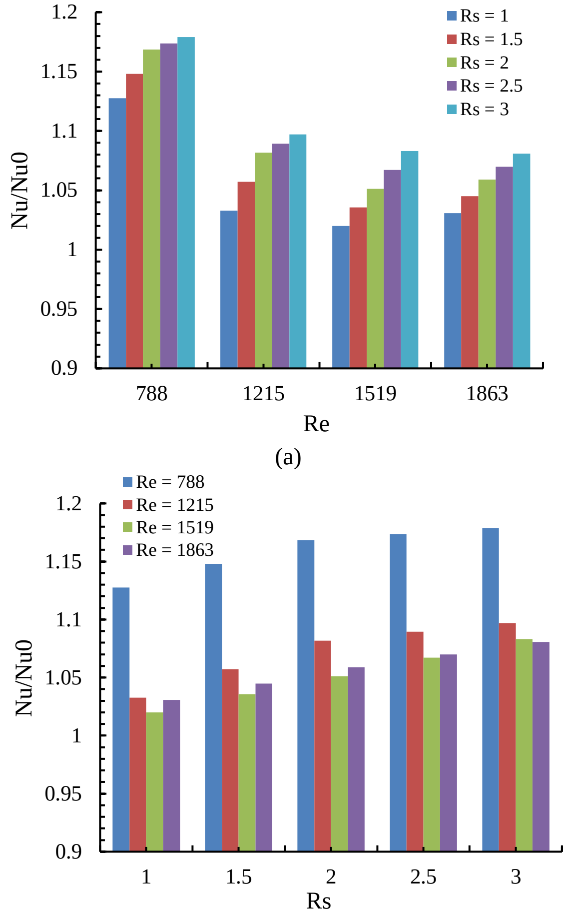


Figure 5. Nu number vs (a) Re number and (b) Rs



(b)

Figure 6. Nu/Nu_0 number vs A) Re number and B) Rs

very important parameters in the evaluation of heat transfer enhancement methods. Due to the simultaneous consideration of the influence of both friction factor and Nu number PEC is very useful in evaluating heat transfer enhancement techniques. Analyzing this parameter allows the engineers to choose the best heat transfer enhancement technique for a certain application. The variation of PEC vs dimensionless rotational speed is presented in Figure 7. As is shown by increment of the Rs the PEC decreases. This point denotes that the increment of friction factor is dominant to the increment of Nu number. The best amounts for the PEC is achieved for Re= 788 and Rs= 1. This point shows that at the lower Re numbers and rotational speeds the presented technique could be more effective than other cases. As shown in Figure 6, it could be seen that the thermal efficiency of the system in Reynolds number of 1863 against the trend. As mentioned before, this values of Reynolds number falls in the transient regime. In this regime the Nusselt number increases to the turbulent nature of the flow. Consequently, in this Reynolds number the thermal efficiency factor finds improved values.

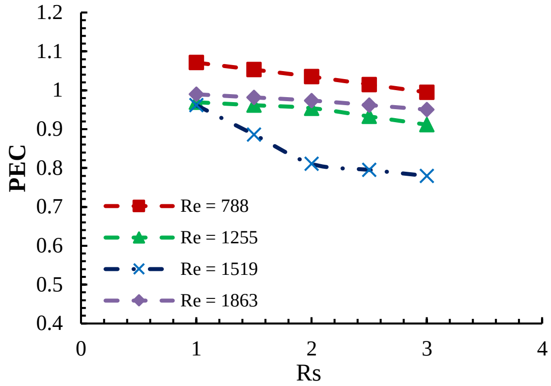


Figure 7. Variation of PEC vs Rs

4. 4. The Cost Per Benefit Ratio Factor The C.B.R factor is the ratio of variation percentage of pressure drop per variation percentage of Nu number. This parameter is another important parameter which could be notified as an assessment criterion for heat transfer enhancement techniques. As is presented in Figure 8 the C.B.R factor increases with the increment of dimensionless rotational speed. The best values of C.B.R factor (lowest amounts) are related to lower amounts of Rs denoting this point that the rotational shaft could be more effective in low rotational speeds. However, at the low rotational speeds the C.B.R values are more than unite. This points out that the pressure drop increment is always dominant to the increment of Nu number.

5. CONCLUSIONS

Present study experimentally evaluated the influence of a rotational circular shaft on thermal behavior of a circular tube. The circular tube was put under fixed heat flux. The heat flux was produced via heater wire which was evenly twisted around the circular tube. The key results of this study are summarized as following.

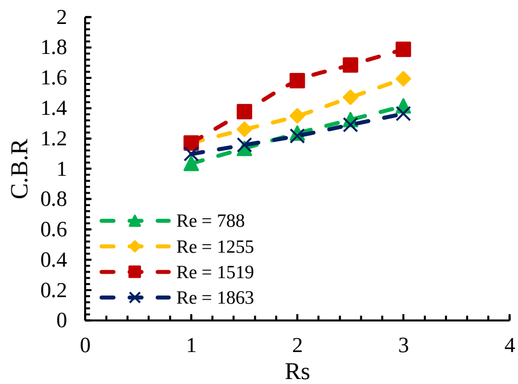


Figure 8. Variation of C.B.R vs Rs

The rotational shaft increases the Nu number.
The maximum augmentation was around 19%.
The rotational shaft could significantly increase the pressure drop and friction factor.
The maximum increment of 78% was achieved for friction factor.
From the view point of both PEC and C.B.R factors the use of rotational shaft to increase the thermal performance of the heat exchanger could be more efficient at low Re numbers and low dimensionless rotational speed. However, by the increment of Re number and reaching to the transient regime of the flow the behavior of the system differs and gets better.

6. REFERENCES

1. Kumar, B., Srivastava, G.P., Kumar, M. and Patil, A.K., "A review of heat transfer and fluid flow mechanism in heat exchanger tube with inserts", ***Chemical Engineering and Processing-Process Intensification***, Vol. 123, (2018), 126-137. https://doi.org/10.1016/j.cep.2017.11.007
2. Kareem, Z.S., Jaafar, M.M., Lazim, T.M., Abdullah, S. and Abdulwahid, A.F., "Passive heat transfer enhancement review in corrugation", ***Experimental Thermal and Fluid Science***, Vol. 68, (2015), 22-38. https://doi.org/10.1016/j.expthermflusci.2015.04.012
3. Gupta, N.K., Tiwari, A.K. and Ghosh, S.K., "Heat transfer mechanisms in heat pipes using nanofluids–a review", ***Experimental Thermal and Fluid Science***, Vol. 90, (2018), 84-100. https://doi.org/10.1016/j.matpr.2017.01.119
4. Gugulothu, R., Somanchi, N.S., Reddy, K.V.K. and Akkiraju, K., "A review on enhancement of heat transfer in heat exchanger with different inserts", ***Materials Today: Proceedings***, Vol. 4, No. 2, (2017), 1045-1050. https://doi.org/10.1016/j.matpr.2017.01.118
5. White, A., "Flow of a fluid in an axially rotating pipe", ***Journal of Mechanical Engineering Science***, Vol. 6, No. 1, (1964), 47-52. https://doi.org/10.1243/JMES_JOUR_1964_006_010_02
6. Suppiah Singaram, S., Lodha, H. and Jachuck, R.J., "Experimental investigation of continuous single-phase rimming flow in a horizontal rotating cylinder", ***AIChE Journal***, Vol. 60, No. 11, (2014), 3939-3950. DOI:10.1002/aic.14569
7. Chatterjee, S., Sugilal, G. and Prabhu, S., "Flow transitions in a partially filled rotating inclined pipe with continuous flow", ***Experimental Thermal and Fluid Science***, Vol. 83, (2017), 47-56. https://doi.org/10.1016/j.expthermflusci.2016.12.007
8. Krivosheev, B., Kukharskii, M. and Portnov, V., "Heat transfer in the evaporator section of a rotating heat pipe at low rotational speeds", ***Inzhenerno Fizicheskii Zhurnal***, Vol. 37, (1979), 27-34. https://doi.org/10.1016/j.ijheatmasstransfer.2007.07.020
9. Chatterjee, S., Sugilal, G. and Prabhu, S., "Heat transfer in a partially filled rotating pipe with single phase flow", ***International Journal of Thermal Sciences***, Vol. 125, (2018), 132-141. DOI:10.1016/J.IJTHERMALSCI.2017.11.024
10. Hussain, S.H. and Hussein, A.K., "Mixed convection heat transfer in a differentially heated square enclosure with a conductive rotating circular cylinder at different vertical locations", ***International Communications in Heat and Mass Transfer***, Vol. 38, No. 2, (2011), 263-274. https://doi.org/10.1016/j.icheatmasstransfer.2010.12.006
11. Park, Y.G., Yoon, H.S. and Ha, M.Y., "Natural convection in square enclosure with hot and cold cylinders at different vertical locations", ***International Journal of Heat and Mass Transfer***, Vol. 55, No. 25-26, (2012), 7911-7925. https://doi.org/10.1016/j.ijheatmasstransfer.2012.08.012
12. Yoon, H.S., Park, Y.G. and Jung, J.H., "Natural convection in a square enclosure with differentially heated two horizontal cylinders", ***Numerical Heat Transfer, Part A: Applications***, Vol. 65, No. 4, (2014), 302-326. https://doi.org/10.1080/10407782.2013.831679
13. Kareem, A.K. and Gao, S., "Mixed convection heat transfer of turbulent flow in a three-dimensional lid-driven cavity with a rotating cylinder", ***International Journal of Heat and Mass Transfer***, Vol. 112, (2017), 185-200. https://doi.org/10.1016/j.ijheatmasstransfer.2017.04.118
14. Selimefendigil, F. and Öztop, H.F., "Numerical study and identification of cooling of heated blocks in pulsating channel flow with a rotating cylinder", ***International Journal of Thermal Sciences***, Vol. 79, (2014), 132-145. DOI:10.1016/j.ijthermalsci.2014.01.010
15. Alawadhi, E.M., "Numerical simulation of fluid flow past an oscillating triangular cylinder in a channel", ***Journal of Fluids Engineering***, Vol. 135, No. 4, (2013). https://doi.org/10.1115/1.4023654
16. Sutherland, B. and Linden, P., "Internal wave excitation by a vertically oscillating elliptical cylinder", ***Physics of Fluids***, Vol. 14, No. 2, (2002), 721-731. https://doi.org/10.1063/1.1430438
17. Pourgholam, M., Izadpanah, E., Motamedi, R. and Habibi, S.E., "Convective heat transfer enhancement in a parallel plate channel by means of rotating or oscillating blade in the angular direction", ***Applied Thermal Engineering***, Vol. 78, (2015), 248-257. https://doi.org/10.1016/j.applthermaleng.2014.12.057
18. Moffat, R.J., "Describing the uncertainties in experimental results", ***Experimental Thermal and Fluid Science***, Vol. 1, No. 1, (1988), 3-17.
19. Rezaei, R.A., Jafarmadar, S. and Khorasani, S., "Presentation of frictional behavior of micro helical tubes with various geometries and related empirical correlation; an experimental study", ***International Journal of Thermal Sciences***, Vol. 140, (2019), 377-387. https://doi.org/10.1016/j.ijthermalsci.2019.03.011
20. Khorasani, S., Jafarmadar, S., Pourhedayat, S., Abdollahi, M.A.A. and Heydarpour, A., "Experimental investigations on the effect of geometrical properties of helical wire turbulators on thermal performance of a helically coiled tube", ***Applied Thermal Engineering***, Vol. 147, (2019), 983-990. https://doi.org/10.1016/j.applthermaleng.2018.09.092
21. Sun, M. and Zeng, M., "Investigation on turbulent flow and heat transfer characteristics and technical economy of corrugated tube", ***Applied Thermal Engineering***, Vol. 129, (2018), 1-11. https://doi.org/https://doi.org/10.1016/j.applthermaleng.2017.09.136
22. Promvonge, P., "Thermal augmentation in circular tube with twisted tape and wire coil turbulators", ***Energy Conversion and Management***, Vol. 49, No. 11, (2008), 2949-2955. DOI: 10.1016/j.enconman.2008.06.022

Persian Abstract

چکیده

روشهای فعال و غیرفعال دو مکانیسم اصلی بهبود انتقال حرارت هستند. روشهای فعال از نیروهای خارجی برای بهبود انتقال گرما استفاده می کنند. این تحقیق رفتار حرارتی و اصطکاکی یک لوله دایره ای حاوی شافت چرخشی را ارزیابی می کند. شار حرارتی ثابت بر روی لوله دایره ای اعمال شد. دمای ورودی و خروجی سیال و همچنین دمای دیواره لوله ها برای محاسبه ضریب انتقال کلاه اندازه گیری شد. تعداد Re (رینولدز) بین 800-2000 بود. همچنین ، سرعت چرخش بدون بعد (Rs) مقادیر 1،1.5 ، 2 ، 2.5 و 3 را نشان داد. نتایج نشان داد که شافت چرخشی می تواند تعداد Nu را افزایش دهد. تا٪ 18 همچنین ، نتایج نشان داد که شافت چرخشی می تواند به طور قابل توجهی افت فشار و ضریب اصطکاک را افزایش دهد. حداکثر افزایش 78٪ برای فاکتور اصطکاک حاصل شد. مشخص شد که استفاده از شافت چرخشی می تواند در تعداد کم Re و سرعت چرخش بدون بعد کم کاراتر باشد. همچنین ، مشخص شد که با افزایش تعداد رینولدز و قرار گرفتن در رژیم گذرا ، کارایی سیستم بهبود می یابد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Non-linear Axial Vibrations of Composite Drill Strings Considering Interaction of Roller Cone Bit and Polycrystalline Diamond Compact Bit with Rock

M. Mohammadzadeh[a], M. Arbabtafti*[a], M. Shahgholi[a], J. Yang[b]

[a] *Faculty of Mechanical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran*
[b] *Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL, Canada*

PAPER INFO

Paper history:
Received 10 November 2020
Received in revised form 18 December 2020
Accepted 03 February 2021

Keywords:
Bit-rock Interaction
Composite Drill Strings
Finite Element
Non-linear Axial Vibrations

ABSTRACT

Due to the world increasing energy demands, optimizing the drilling system parameters such as the weight on bit (WOB), and the structure of drill string and bit, also vibrations and dynamic behavior of drill strings are of significant interest to researchers and energy industries. Specially, to overcome limitations of drilling operations in oil and gas industry, composite drill strings as high-tech devices are under development. In this research, the fully coupled non-linear axial vibrations of composite drill strings due to the interaction of two common bits namely; Roller Cone (RC) and Polycrystalline Diamond Compact (PDC) with rock, considering the major non-linear terms, the drill string-wellbore contact, the different weight on bit (WOB) and the different composite configurations using the finite element method (FEM) and the Lagrangian approach were studied. This study proved that the different configurations of composite drill strings showed specific dynamic behavior at different conditions. Therefore, composite drill string can be designed for particular purposes. Also, the results imply the remarkable effects of weight on bit (WOB) and type of bits on the axial vibrations of composite drill strings.

***doi**: 10.5829/ije.2021.34.04a.28*

1. INTRODUCTION

A drill string is an ultra-slender rotor, which is suspended by a hoisting system on the surface, to connect the rotary drive system at the top to the bit at the bottom to dig through the rock. The drill string is put lower through the well to generate the compressional load applied on the bit F_{bit} , which is generally equal to the value of weight on bit (WOB). The upper and lower fragments of the drill string are the drill pipes and drill collars, respectively (see Figure 1).

The basic experimental and analytical analysis of vibrations of drill stings started in past decades [1-3]. Yigit and Christoforou [4] studied the coupled axial and transverse vibrations of oilwell drillstrings. Optimizing the drilling parameters such as weight on bit, rate of penetration (ROP) and structure of drill string and bit are of great interest to researchers and energy industries, academically and practically [5-8]. The effects of downhole assembly and polycrystalline diamond compact (PDC) bit geometry on stability of drillstrings were studied by Elsayed et al. [9]. Jansen [10] analyzed the dynamic behavior of drill string considering the drilling fluid and drill string-wellbore contact. Jogi et al. [11] presented four programs to analyze vibrations. Spanos et al. [12] studied the drill string-wellbore contact using a nonlinear finite element dynamic model. Khulief and Al-Naser [13] analyzed dynamic of a finite element model of drill string. The system responses with the reduced-order and full-order models were compared. Ghasemloonia et al. [14] studied the coupled non-linear axial-transverse vibration of a drill string using the Bypassing PDE's method with the expanded Galerkin's method. Nowadays, the study on optimization of materials to improve their role in practical aspects are increasing [15-17]. In recent years, the use of new torsion shafts [18,19] and composite

*Corresponding Author Institutional Email: *Arbabtafti@sru.ac.ir* (M. Arbabtafti)

Please cite this article as: M. Mohammadzadeh, M. Arbabtafti, M. Shahgholi, J. Yang, Non-linear Axial Vibrations of Composite Drill Strings Considering Interaction of Roller Cone Bit and Polycrystalline Diamond Compact Bit with Rock, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1016-1023

rotors [20, 21] is expanding. Specially, to overcome limitations of drilling operations in oil and gas industry, composite drill strings as high-tech devices are under development.

The composite drill strings can be designed to meet specific requirements for specific applications, e.g. flexible enough to resist bending fatigue, but strong enough to carry high tensional and torsional loads, particularly in the ultra-deep directional drilling or the short radius directional drilling [20-22].

Mohammadzadeh et al. [22] furnished a basic fully coupled non-linear model of composite drill strings for the further development of a more comprehensive model. The approach presented on that work[22] has been adapted in recent study and has been extended to accommodate the important case of non-linear axial vibrations considering bit (RC/PDC)-rock interaction.

This research has aimed to study the fully coupled non-linear axial vibrations of vertical composite drill strings due to the interaction of two common bits namely; Roller-Cone (RC) bits and Polycrystalline Diamond Compact (PDC) with rock, considering the drill string-wellbore contact, the different weight on bit (WOB) and the different composite configurations. Also, the gyroscopic effect, and especially the geometric stiffening effect, the axial-torsional-lateral coupling of vibrations and the major non-linear terms have been taken into account. The full-order non-linear equations of the whole length of drill string including drill pipes and drill collars have been derived by the finite element method and the Lagrangian approach.

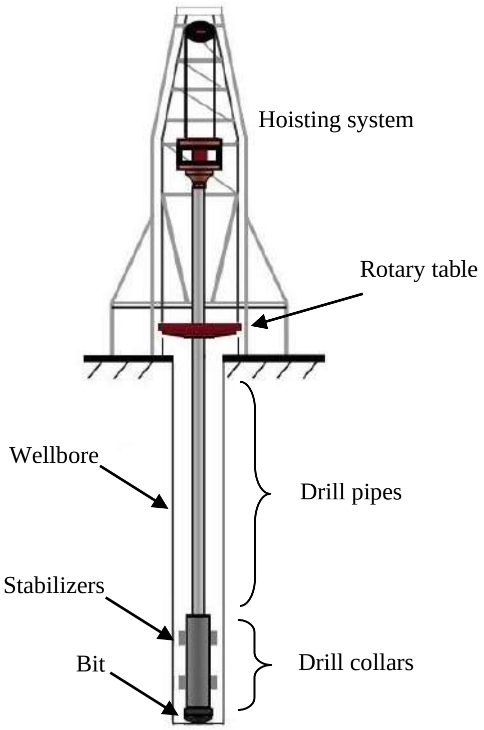


Figure 1. The drilling system

2. PROBLEM FORMULATION

The laminated composite drill string consists of the winded orthotropic layers.

2. 1. Finite Element Discretization

Using a number of two-node elements with six degrees of freedom per node, the element displacement vector {e} can be defined by Equation (1):

$$\{\boldsymbol{e}\}_{12\times1} = [u_1 \;\; v_1 \;\; w_1 \;\; \theta_{x1} \;\; \theta_{y1} \;\; \theta_{z1} \;\; u_2 \;\; v_2 \;\; w_2 \;\; \theta_{x2} \;\; \theta_{y2} \;\; \theta_{z2}]^{tr} \tag{1}$$

Figure 2 shows the translational displacements u, v, w, the torsional displacements θ_x , and the rotational displacements θ_y and θ_z in their directions. The nodal displacement functions can be expressed below:

$$\begin{aligned} u_e(x,t) &= N_u(x)\,\boldsymbol{e}(t) & \theta_{xe}(x,t) &= N_{\theta_x}(x)\,\boldsymbol{e}(t) \\ v_e(x,t) &= N_v(x)\,\boldsymbol{e}(t) & \theta_{ye}(x,t) &= N_{\theta_y}(x)\,\boldsymbol{e}(t) \\ w_e(x,t) &= N_w(x)\,\boldsymbol{e}(t) & \theta_{ze}(x,t) &= N_{\theta_z}(x)\,\boldsymbol{e}(t) \end{aligned} \tag{2}$$

where N is the shape function based on Timoshenko beam theory [23-25].

2. 1. Kinetic Energy of the Composite Drill String

The kinetic energy of an element can be written as follows:

$$T = \frac{1}{2}\int_0^{L_e} \{ m\mathbf{V}^{tr}\mathbf{V} + \boldsymbol{\omega}^{tr}[\mathrm{I}]\boldsymbol{\omega} \}\, dx \tag{3}$$

where m is the mass per unit length, $\mathbf{V}$ is translation velocity vector of the cross section, $\boldsymbol{\omega}$ is angular

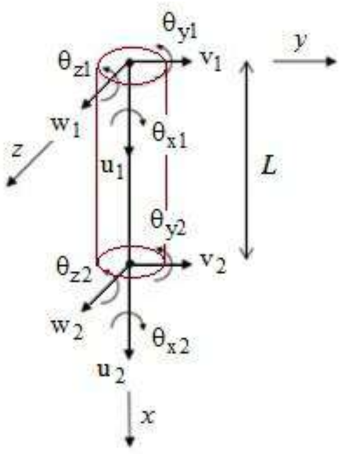


Figure 2. Degrees of freedom of an element

velocity vector in the fixed reference system and [I] is the matrix of mass moment of inertia stated as follows:

$$m = \frac{\pi}{4}\sum_{k=1}^{n}\rho_k (D_{ok}^2 - D_{ik}^2),\quad \mathbf{V} = \begin{pmatrix} \frac{\partial u}{\partial t} & \frac{\partial v}{\partial t} & \frac{\partial w}{\partial t} \end{pmatrix}^{tr}$$

$$\boldsymbol{\omega} = \begin{pmatrix} \omega_x & \omega_y & \omega_z \end{pmatrix}^{tr},\quad [\mathrm{I}] = Diag\,[I_P \quad I_D \quad I_D]$$

Where

$$\omega_x = (\frac{\partial \Phi}{\partial t} + \frac{\partial \theta_x}{\partial t}) - \frac{\partial \theta_z}{\partial t}\theta_y \tag{4}$$

$$\omega_y = \frac{\partial \theta_y}{\partial t}\cos(\Phi + \theta_x) - \frac{\partial \theta_z}{\partial t}\sin(\Phi + \theta_x)$$

$$\omega_z = \frac{\partial \theta_y}{\partial t}\sin(\Phi + \theta_x) + \frac{\partial \theta_z}{\partial t}\cos(\Phi + \theta_x)$$

That I_P and I_D are the polar and diametrical mass moment of inertia. After some algebraic manipulations, the total kinetic energy of an element can be written as follows:

$$\begin{aligned} T = \frac{1}{2}\int_0^{L_e} \Bigg(& \frac{\pi}{4}\sum_{k=1}^{n}\rho_k(D_{ok}^2 - D_{ik}^2)\left(\frac{\partial u}{\partial t}\right)^2 + \frac{\pi}{4}\sum_{k=1}^{n}\rho_k(D_{ok}^2 - D_{ik}^2)\left(\frac{\partial v}{\partial t}\right)^2 \\ & + \frac{\pi}{4}\sum_{k=1}^{n}\rho_k(D_{ok}^2 - D_{ik}^2)\left(\frac{\partial w}{\partial t}\right)^2 + \frac{\pi}{32}\sum_{k=1}^{n}\rho_k(D_{ok}^4 - D_{ik}^4)\left(\frac{\partial(\Phi+\theta_x)}{\partial t}\right)^2 \\ & + \frac{\pi}{64}\sum_{k=1}^{n}\rho_k(D_{ok}^4 - D_{ik}^4)\left(\frac{\partial \theta_y}{\partial t}\right)^2 + \frac{\pi}{64}\sum_{k=1}^{n}\rho_k(D_{ok}^4 - D_{ik}^4)\left(\frac{\partial \theta_z}{\partial t}\right)^2 \\ & - \frac{\pi}{32}\sum_{k=1}^{n}\rho_k(D_{ok}^4 - D_{ik}^4)\left(2\left(\frac{\partial(\Phi+\theta_x)}{\partial t}\right)\left(\frac{\partial \theta_z}{\partial t}\right)\theta_y - \left(\frac{\partial \theta_z}{\partial t}\right)^2\theta_y^2\right)\Bigg) dx \end{aligned} \tag{5}$$

That accounts for the translational inertia, the torsional inertia, the rotational inertia, and the coupling term comprise the non-linear terms and the gyroscopic moments. Also, n is the number of the layers, ρ_k is the material density of the *k-th* layer. As shown in Figure 3, D_{ok} and D_{ik} are the external and internal diameter of the *k-th* layer, respectively.

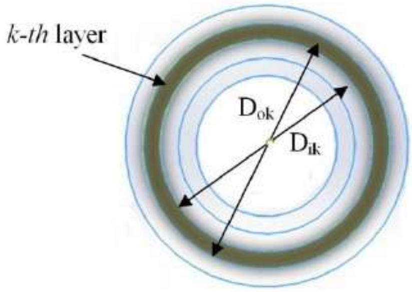


Figure 3. The external and internal diameter of the *k-th* layer

2. 2. Strain Energy of the Composite Drill String

The generalized Hooke's law for fiber-reinforced composite materials in the cylindrical coordinate system can be expressed as follows:

$$\{\sigma\}_{x,\theta,r} = [\bar{C}]\{\varepsilon\}_{x,\theta,r} \tag{6}$$

Where $\{\sigma\}$ and $\{\varepsilon\}$ are the stress and strain fields, and $[\bar{C}]$ is the transformed stiffness matrix of a layer [26].

The strain energy of an element in the cylindrical coordinate system is given below:

$$\begin{aligned} U = \frac{1}{2}\int_V \{ & \sigma_{xx}\varepsilon_{xx} + \sigma_{rr}\varepsilon_{rr} + \sigma_{\theta\theta}\varepsilon_{\theta\theta} \\ & + 2\tau_{xr}\varepsilon_{xr} + 2\tau_{x\theta}\varepsilon_{x\theta} + 2\tau_{r\theta}\varepsilon_{r\theta}\}\, dV \end{aligned} \tag{7}$$

Since the length of drill string is very bigger than the lateral dimensions, therefore:

$\sigma_{rr}, \sigma_{\theta\theta}, \tau_{r\theta} \cong 0$ and $\varepsilon_{rr}, \varepsilon_{\theta\theta}, \varepsilon_{r\theta} \cong 0$, thus:

$$\begin{aligned} U = \frac{1}{2}\int_0^{L_e} \Bigg\{ & \frac{\pi}{4}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^2 - D_{ik}^2)(\frac{\partial u}{\partial x})^2 \\ & + k_s\frac{\pi}{32}\sum_{k=1}^{n}\bar{C}_{66k}(D_{ok}^4 - D_{ik}^4)(\frac{\partial \theta_x}{\partial x})^2 \\ & + \frac{\pi}{64}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^4 - D_{ik}^4)((\frac{\partial \theta_y}{\partial x})^2 \\ & + (\frac{\partial \theta_z}{\partial x})^2) + k_s(\frac{\pi}{8}\sum_{k=1}^{n}\bar{C}_{55k}(D_{ok}^2 - D_{ik}^2) \\ & + \frac{\pi}{8}\sum_{k=1}^{n}\bar{C}_{66k}(D_{ok}^2 - D_{ik}^2))((\frac{\partial v}{\partial x} - \theta_z)^2 + (\frac{\partial w}{\partial x} + \theta_y)^2) \\ & + \frac{\pi}{4}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^2 - D_{ik}^2)((\frac{\partial u}{\partial x})^3 + \frac{\partial u}{\partial x}(\frac{\partial v}{\partial x})^2 + \frac{\partial u}{\partial x}(\frac{\partial w}{\partial x})^2) \\ & + \frac{3\pi}{64}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^4 - D_{ik}^4)(\frac{\partial u}{\partial x}(\frac{\partial \theta_y}{\partial x})^2 + \frac{\partial u}{\partial x}(\frac{\partial \theta_z}{\partial x})^2) \\ & + k_s\frac{\pi}{12}\sum_{k=1}^{n}\bar{C}_{16k}(D_{ok}^3 - D_{ik}^3) \\ & \times((2\frac{\partial \theta_x}{\partial x}\frac{\partial u}{\partial x} - \frac{\partial \theta_y}{\partial x}(\frac{\partial v}{\partial x} - \theta_z) - \frac{\partial \theta_z}{\partial x}(\frac{\partial w}{\partial x} + \theta_y))\Bigg\} dx \end{aligned} \tag{8}$$

Which accounts for the axial, the torsional, the bending and the non-linear shearing deformations. The last three terms represent other major non-linear low-order terms. The lateral stiffness of drill strings varies when they are subjected to high axial loads, this phenomenon is called the geometric stiffening effect [22, 27], the strain energy due to geometric stiffening effect is denoted by U_{gs}:

$$U_{gs}=\frac{1}{2}\int_0^{L_e}[\frac{\pi}{4}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^2-D_{ik}^2)$$
$$\times((\frac{\partial u}{\partial x})^3+(\frac{\partial u}{\partial x})(\frac{\partial v}{\partial x})^2+(\frac{\partial u}{\partial x})(\frac{\partial w}{\partial x})^2) \quad (9)$$
$$+\frac{3\pi}{64}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^4-D_{ik}^4)((\frac{\partial u}{\partial x})(\frac{\partial \theta_y}{\partial x})^2+(\frac{\partial u}{\partial x})(\frac{\partial \theta_z}{\partial x})^2)]dx \quad (9)$$

where, the term $\frac{\pi}{4}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^2-D_{ik}^2)$ is the net axial force $F_A(x)$ [27], then:

$$U_{gs}=\frac{1}{2}\int_0^{L_e}[F_A(x)((\frac{\partial u}{\partial x})^3+(\frac{\partial u}{\partial x})(\frac{\partial v}{\partial x})^2+(\frac{\partial u}{\partial x})(\frac{\partial w}{\partial x})^2) \quad (10)$$
$$+\frac{3\pi}{64}\sum_{k=1}^{n}\bar{C}_{11k}(D_{ok}^4-D_{ik}^4)((\frac{\partial u}{\partial x})(\frac{\partial \theta_y}{\partial x})^2+(\frac{\partial u}{\partial x})(\frac{\partial \theta_z}{\partial x})^2)]dx \quad (10)$$

The net axial force $F_A(x)$ in a point is the resultant of $W(x)$ the downward drill string weight and F_{bit} the upward reaction force between the bit and the bottom of the well formation. The upper parts of the drill string operate under axial tension, while the upward reaction force causes the lower parts operate under axial compression. There is a neutral point in drill strings, in which $F_A(x)$=0. So, $F_A(x)$ and U_{gs} must be defined for compression and tension fields (Figure 4).

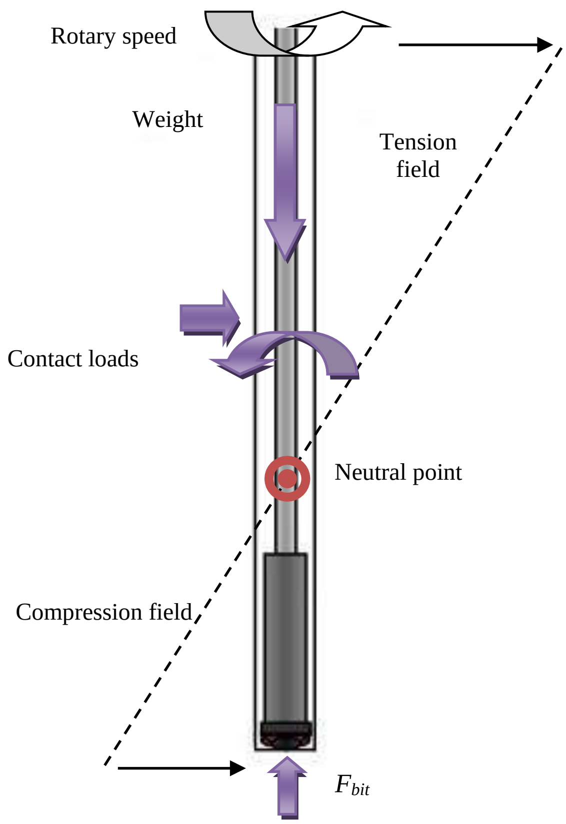


Figure 4. The general scheme of the loads

2. 3. Drill String-wellbore Contact The contact loads between the drill string and wellbore at each point can be modeled by F_n the concentrated forces along normal and F_t tangential directions and M_f the frictional concentrated torque, so:

$$F_n=\begin{cases}0 & \text{for } (\gamma \le g)\\ -k_c(\gamma-g) & \text{for } (\gamma > g)\end{cases},$$
$$F_t=\mu F_n\, sign(\frac{\partial \Phi}{\partial t}), \quad \text{and} \quad M_f=0.5\,D_o\,F_t \quad (11)$$

Where $\gamma=\sqrt{v^2+w^2}$, $g=0.5(\mathrm{D_{ch}}-\mathrm{D_o})$

γ is the radial displacement of the drill string, g is the gap between the surface of the drill string and the wellbore surface, k_c is the stiffness of contact to simulate normal force F_n as an elastic force, μ is the frictional coefficient, D_{ch} and is the wellbore diameter and D_o is the external diameter of drill string.

2. 4. Bit-rock Interaction The main process of drilling in oil and gas industries is the creation of borehole by a rock-cutting tool called a bit. Traditional bits consist of a steel body equipped with three rotating conical cylinder with steel teeth that crush the rock, Roller-Cone (RC) bits. Modern bits often consist of a steel body without rotating parts, covered with artificial diamond cutters that shear the rock, Polycrystalline Diamond Compact (PDC) bits (Figures 5).

The frequency of contact force between PDC bit and rock is equal to the angular velocity of rotary table, while the frequency of contact force between RC bit and rock is three times the angular velocity of rotary table [9]. So, the two continous forces, which are applied in the axial direction at the end of the bit, have been assumed to simulate the interaction of these bits and rock, respectively. The forces are expressed as follow:

$$\mathrm{F_{RC-bit}}=\mathrm{WOB}\sin(3\Omega t)$$
$$\mathrm{F_{PDC-bit}}=\mathrm{WOB}\sin(\Omega t) \quad (12)$$

2. 5. Dynamic Equation of Motion The dynamic equation of motion can be derived using the Lagrange's equation, which can be given below:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)-\frac{\partial L}{\partial q}=Q \quad (13)$$

where $L=(T-U)$ is the Lagrangian function, q and Q are the vectors of generalized coordinates and


Figure 5 (a) RC bit

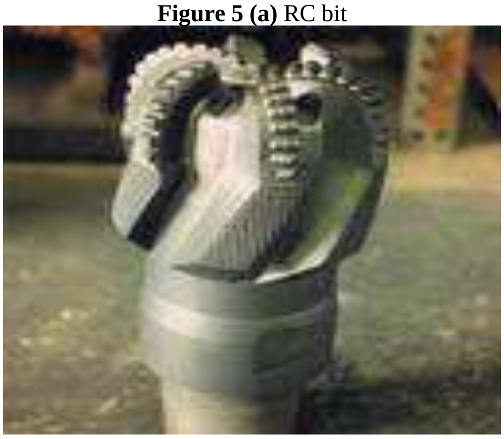


Figure 5 (b) PDC bit

generalized forces, respectively. The expressions of the component terms of Equation (1) have been integrated into computational scheme to obtain the results.

3. SIMULATION AND RESULTS ANALYSIS

The full-order non-linear dynamic equations of composite drill string were solved using a computational plan in MATLAB. The drill string was discretized into 240 elements to achieve results convergence. Table 1 shows the two 16-orthotropic-layer cases with different configurations to show major trends. Also, Table 2 displays the used data in simulation.

It is assumed that the lateral translations and the lateral rotations are zero at the top. The lateral translations at the bit are also zero. The vertical drill string at the top node is subjected to the constant angular velocity about the x-axis and then WOB is applied. The general scheme of loads has been displayed in Figure 4.

TABLE 1. The under investigation cases

Case	1	2
Configuration	$[\pm75]_8$	$[\pm15]_8$

TABLE 2. The used data in simulation

Drill string length = 1200 m	E_1=141.343 GPa
Drill pipes length = 1000 m	E_2= E_3=9.563 GPa
Drill pipes OD/ID = 0.1524/0.127 m	G_{12}= G_{13}=4.55 GPa
Drill collars length =200 m	G_{23}=2.85 Gpa
Drill collars OD/ID = 0.2324/0.1062 m	ν_{12}= ν_{13}=0.28
Wellbore diameter = 0.3524 m	ν_{23}=0.517
Contact stiffness = 10e8 N/m	ρ= 3930 kg/m^3
Frictional coefficient = 5e-4	
Constant angular velocity = 6 Rad/s	
The first stabilizer location = 50 m above the bottom,	
The second stabilizer location = 150 m above the bottom	

The WOB values have been assumed 10 and 30% of the drill collars weight. Figures 6-13 show the axial vibrations of end point of drill collar, which is located at 200 m above the bit.

Figures 6 and 7 show the axial deflections, when WOB is 10% of the drill collars weight for PDC bits in cases 1 and 2, respectively.

The axial deflections, when WOB is 10% of the drill collars weight for RC bits in cases 1 and 2 are displayed in Figures 8 and 9, respectively.

When the ply angle decreases, due to increase of axial stiffness, firstly; the maximum amplitudes of responses decrease; secondly, the capability of drill string to transmit the axial vibrations increases, and the drill string become more sensitive to excitations, so the drill string experience the denser axial vibrations at a point above the bit, as shown in Figures 6 and 7.

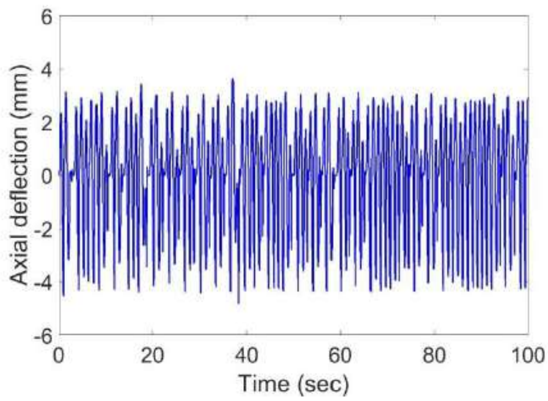


Figure 6. Axial deflection at 200 m above the bit, WOB is 10% of the drill collars weight, PDC bit, (Case 1)

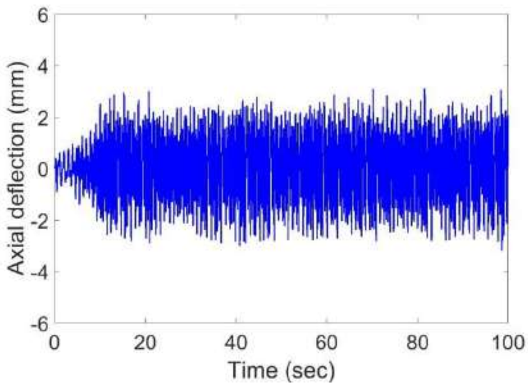


Figure 7. Axial deflection at 200 m above the bit, WOB is 10% of the drill collars weight, PDC bit, (Case 2)

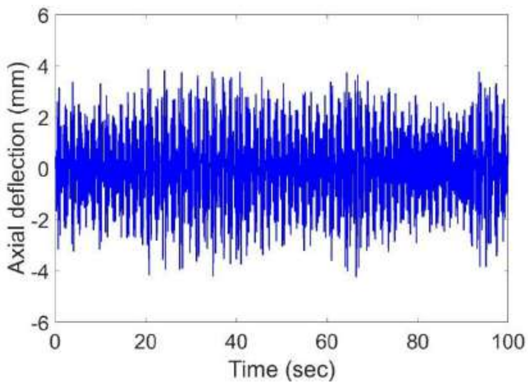


Figure 8. Axial deflection at 200 m above the bit, WOB is 10% of the drill collars weight, RC bit, (Case 1)

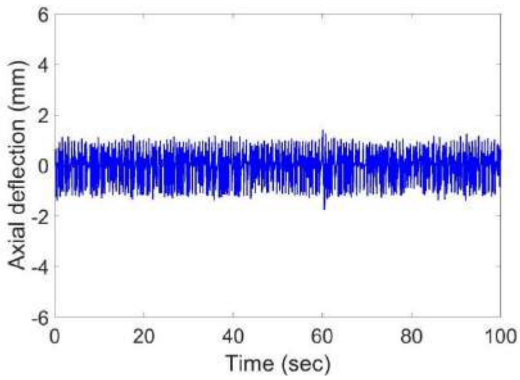


Figure 9. Axial deflection at 200 m above the bit, WOB is 10% of the drill collars weight, RC bit, (Case 2)

As displayed in Figures 8 and 9, Case 2 shows the denser axial vibrations with less amplitudes at a point above the bit. Also, the comparison between these figures with Figures 6 and 7 show that using RC bit intensify the density of axial vibrations.

Figures 10 and 11 show the axial deflections when WOB is 30% of the drill collars weight for PDC bits in cases 1 and 2, respectively.

As shown in Figures 10 and 11, the increase of WOB leads to more axial vibrations. Although due to capturing the response of a point that located in thick part of drill string, this result is not remarkable.

The axial deflections, when WOB is 30% of the drill collars weight for RC bits in cases 1 and 2 are displayed in Figures 12 and 13, respectively.

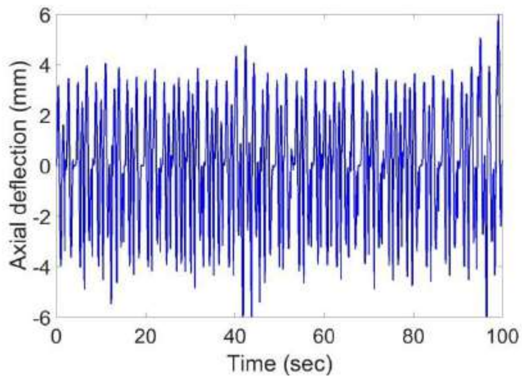


Figure 10. Axial deflection at 200 m above the bit, WOB is 30% of the drill collars weight, PDC bit, (Case 1)

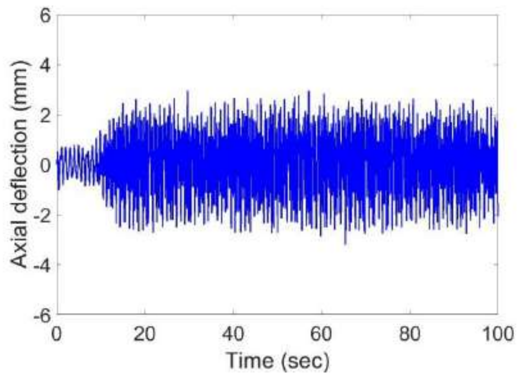


Figure 11. Axial deflection at 200 m above the bit, WOB is 30% of the drill collars weight, PDC bit, (Case 2)

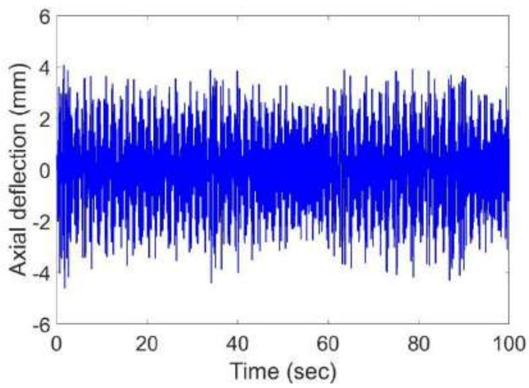


Figure 12. Axial deflection at 200 m above the bit, WOB is 30% of the drill collars weight, RC bit, (Case 1)

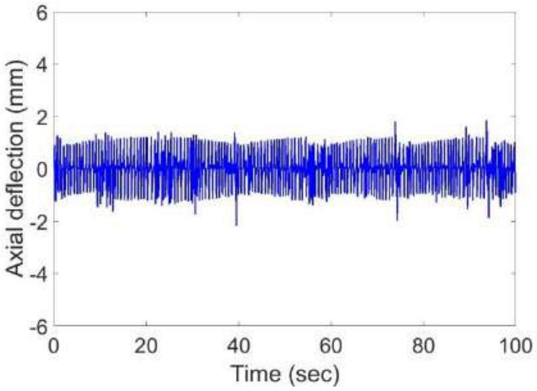


Figure 13. Axial deflection at 200 m above the bit, WOB is 30% of the drill collars weight, RC bit, (Case 2).

As shown in Figures 12 and 13, when the RC bits are used, the axial responses in a point above the bit are denser than PDC bits.

4. CONCLUSION

In this study, the fully coupled non-linear axial vibrations of composite drill strings due to the interaction of two common bits namely; Roller-Cone (RC) bits and Polycrystalline Diamond Compact (PDC) with rock, considering the drill string-wellbore contact, the different weight on bit (WOB) and the different composite configurations were investigated. The gyroscopic effect, and especially the geometric stiffening effect, the axial-torsional-lateral coupling of vibrations and the major non-linear terms have been taken into account. The full-order non-linear equations of the whole length of drill string including drill pipes and drill collars were derived by the finite element method and the Lagrangian approach.

This study proved that the different configurations of composite drill strings showed specific dynamic behavior in different conditions; therefore, composite drill string can be designed for particular purposes. Also, the results imply the remarkable effects of weight on bit and type of bits on the axial vibrations of composite drill strings.

As discussed, when the ply angle decreases, due to increase of axial stiffness, firstly; the maximum amplitudes of responses decrease; secondly, the capability of drill string to transmit the axial vibrations increases, and the drill string become more sensitive to excitations, so the drill string experience the denser axial vibrations at the point above the bit. Also, using RC bit intensifies the density of axial vibrations and the increase of weight on bit leads to more axial vibrations.

5. REFERENCES

1. Finnie, I. and Bailey, J.J., "An experimental study of drillstring vibration", ***Journal of Engineering for Industry***, ASME Transaction Vol. 82, No. 2, (1960), 129-135, doi: 10.1115/1.3663020.
2. Bailey, J.J. and Finnie, I., "An analytical study of drillstring vibration", ***Journal of Engineering for Industry***, ASME Transaction Vol. 82, No. 2, (1960), 122-128, doi: 10.1115/1.3663017.
3. Wolf, S.F., Zacksenhouse, M. and Arian, A., "Field measurements of downhole drillstring vibrations", Proceedings of the Annual Technical Conference & Exhibition, Las Vegas, (1985), 32-42, doi: 10.2118/14330-MS.
4. Yigit, A.S. and Christoforou, A.P., "Coupled axial and transverse vibrations of oilwell drillstrings", ***Journal of Sound and Vibration***, Vol. 195, No. 4, (1996), 617-627, doi: 10.1006/jsvi.1996.0450.
5. Wu, X., Paez, L.C., Partin, U.T. and Agnihotri, M., "Decoupling stick-slip and whirl to achieve breakthrough in drilling performance", Proceedings of the IADC/SPE Drilling Conference and Exhibition, Society of Petroleum Engineers (2010), doi: 10.2118/128767-MS.
6. Bailey J.R. and Remmert, S.M., "Managing drilling vibrations through BHA design optimization", ***Drilling and Completion***, Vol. 25, No. 40, (2010), 458-471, doi: 10.2118/139426-PA.
7. Dvoynikov, M., Syzrant sevb, V. and Syzrant sevab, K., "Designing a high resistant, high-torque downhole drilling motor", ***International Journal of Engineering***, ***Transactions A: Basics***, Vol. 30, No. 10, (2017), 1615-1621, doi: 10.5829/ije.2017.30.10a.24.
8. Wang, H., Guan, Z.C., Shi, Y.C., Liu, Y.W. and Liang D.Y., " Drilling Trajectory Prediction Model for Push-the bit Rotary Steerable Bottom Hole Assembly", ***International Journal of Engineering, Transactions B: Applications***, Vol. 30, No. 11, (2017), 1800-1806, doi: 10.5829/ije.2017.30.11b.23.
9. Elsayed, M.A., Dareing, D.W. and Dupuy, C.A. "Effect of downhole assembly and polycrystalline diamond compact (PDC) bit geometry on stability of drillstrings", ***Journal of Energy Resources Technology***, Vol. 119, (1997), 159-163, doi: 10.1115/1.2794984.
10. Jansen, J.D., "Nonlinear rotor dynamics as applied to oilwell drillstring vibrations", ***Journal of Sound and Vibration***, Vol. 147 No. 1, (1991), 115-135, doi: 10.1016/0022-460X(91)90687-F.
11. Jogi, P.N., Macpherson, J.D. and Neubert, M., "Field verification of model-derived natural frequencies of a drill string", ***Journal of Energy Resources Technology***, Vol. 124, No. 3, (2002), 154-162, doi: 10.1115/1.1486018.
12. Spanos, P.D., Chevallier, A.M. and Politis, N.P., "Nonlinear stochastic drill-string vibrations", ***Journal of Applied Mechanics***, ASME Transactions, Vol. 124, (2002), 512-518, doi: 10.1115/1.1502669.
13. Khulief, Y.A. and Al-Naser, H., "Finite element dynamic analysis of drillstrings", ***Finite Element in Analysis and Design*** Vol. 41, (2005), 1270–1288, doi: 10.1016/j.finel.2005.02.003.
14. Ghasemloonia, A., Rideout, D.G. and Butt, S.D., "Analysis of multi-mode nonlinear coupled axial-transverse drillstring vibration in vibration assisted rotary drilling", ***Journal of Petroleum Science and Engineering***, Vol. 116, (2014), 36-49, doi: 10.1016/j.petrol.2014.02.014.
15. Trang, G.T.T., Linh, N.H., Linh, N.T.T. and Kien, P.H., "The study of dynamics heterogeneity in SiO_2 liquid", ***High Tech and Innovation Journal***, Vol. 1, No. 1, (2020), 1-7, doi: 10.28991/HIJ-2020-01-01-01.

16. Fazelabdolabadi, B. and Golestan, M.H. "Towards bayesian quantification of permeability in micro-scale porous structures – The database of micro networks", ***High Tech and Innovation Journal***, Vol. 1, No. 4, (2020), 148-160, doi: 10.28991/HIJ-2020-01-04-02.
17. Battistelli, D., Ferreira, D.P., Costa, S., Santulli C. and Fangueiro, R. "Conductive thermoplastic starch (TPS) composite filled with waste iron filings", ***Emerging Science Journal***, Vol. 4, No. 3, (2020), 136-147, doi: 10.28991/esj-2020-01218.
18. Furch, J., Nguyen, Q.H. "Lifetime test of tracked vehicle torsion bars using Monte Carlo method", ***Emerging Science Journal***, Vol. 4, No. 5, (2020), 376-389, doi: 10.28991/esj-2020-01238.
19. Ha, K., "Innovative blade trailing edge flap design concept using flexible torsion bar and worm drive", ***High Tech and Innovation Journal***, Vol. 1, No. 3, (2020), 101-106, doi: 10.28991/HIJ-2020-01-03-01.
20. J.C. Leslie, J.T. Heard, L. Truong, "Advances in composite drilling components lead to evaluation for critical E&P applications", ***Advanced Composite Products and Technology***, (2007).
21. Dong, G. and Chen, P. "A review of the evaluation, control, and application technologies for drill string vibrations and shocks in oil and gas well", ***Shock and Vibration***, (2016), Article ID 7418635, doi: 10.1155/2016/7418635.
22. Mohammadzadeh, M., Shahgholi, M., Arbabtafti, M. and Yang, J., "Vibration analysis of the fully coupled nonlinear finite element model of composite drill strings", ***Archive of Applied Mechanics***, Vol. 90, (2020), 1373–1398, doi: 10.1007/s00419-020-01673-8.
23. Bazoune, A., Khulief, Y.A. and Stephen, N.G., "Shape functions of three-dimensional Timoshenko beam element", ***Journal of Sound and Vibration***, Vol. 259, No. 2, (2003), 473-480, doi: 10.1006/jsvi.2002.5122.
24. Su, Y.Y. and Gao, K.L., "Analytical model for adhesively bonded composite panel-flange joints based on the Timoshenko beam theory", ***Composite Structures***, Vol. 107, (2014), 112-118, doi: 10.1016/j.compstruct.2013.07.018.
25. Khalkhali, A., Narimanzadeh, N., Khakshournia, S. and Amiri, S., "Optimal design of sandwich panels using multi-objective genetic algorithm and finite element method", ***International Journal of Engineering-Transactions C: Aspects***, Vol. 27, No.3, (2013), 395-402, doi: 10.5829/idosi.ije.2014.27.03c.06.
26. Jones, R.M., "Mechanics of Composite Materials", 2nd Ed, Taylor & Francis Inc. (1999).
27. Likens, P., Barbera, F. and Baddeley, V., "Mathematical modeling of spinning elastic bodies for modal analysis", ***Journal of American Institute of Aeronautics and Astronautics***, Vol. 11, No. 9, (1973), 1251-1258, doi: doi.org/10.2514/3.6906.

Persian Abstract

چکیده

رشته حفاری به عنوان یک تجهیز استراتژیک در صنایع نفت و گاز شناخته شده و همواره مورد توجه مراکز تحقیقاتی و صنایع بوده است. همراه با توسعه علم کامپوزیت و مزایای فراوان آنها، تکنولوژی رشته حفاری کامپوزیت به سرعت در حال توسعه می باشد. هدف از این تحقیق تحلیل ارتعاشات کوپل شده غیر خطی محوری رشته حفاری کامپوزیت با درنظر گرفتن تماس رشته حفاری – دیواره چاه و بخصوص تماس مته حفاری – سازه کف چاه می باشد. دو نوع متدوال مته حفاری با عنوان مته سه کاجه (RC) و مته الماسه (PDC) در نظر گرفته شده اند. رشته حفاری کامپوزیت متشکل از لایه های اورتوتروپ در نظر گرفته شده و برای بدست آوردن نتایج، معادلات دینامیکی غیر خطی مرتبه کامل حاکم بر تمام طول رشته حفاری کامپوزیت عمودی با استفاده از روش انرژی و معادلات لاگرانژ و به کمک روش المان محدود استخراج و حل می گردند. کوپل انواع ارتعاشات محوری- جانبی- پیچشی در معادلات غیر خطی، اثر ژیروسکوپی، ترمهای غیر خطی اصلی و اثر سختی هندسی ناشی از تقابل مته و نیروی وزن رشته حفاری در نظر گرفته شده اند. ارتعاشات کوپل شده غیر خطی محوری رشته حفاری کامپوزیت ناشی از تقابل دو نوع مته حفاری و سازه کف چاه، با در نظر گرفتن وزن روی مته (WOB) متفاوت برای دو رشته حفاری با ساختار کامپوزیتی متفاوت، که توانایی نشان دادن روند کلی تغییرات را دارند، مورد مطالعه قرار گرفته است.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 1024-1034

International Journal of Engineering

Journal Homepage: www.ije.ir

Control of Steel Detachment and Metal Flow on Aluminum-steel Friction Stir Welding of Thin Joints

E. A. Torres[a], J. Graciano-Uribe[b], T. F. A. Santos[c]

[a] *Department of Mechanical Engineering, Research Group - GEA, Universidad de Antioquia, Medellín, Colombia*
[b] *Department of Mechatronics Engineering, Research Group - MATyER, Instituto Tecnológico Metropolitano, Medellín, Colombia*
[c] *Department of Mechanical Engineering, Universidade Federal de Pernambuco, Research group- SOLDAMAT, Recife, Brazil*

PAPER INFO

Paper history:
Received 10 November 2019
Received in revised form 23 December 2019
Accepted 04 January 2020

Keywords:
Dissimilar Joints
Offset Effect
Steel Fragments Formation
Restriction to Metal Flow

ABSTRACT

In the last thirty years, the friction stir welding (FSW) process has achieved significant importance due to the satisfactory results derived from severe deformation and low heat input during the welded joint production. These elements have been considered to implement the FSW in different welded systems, including aluminum-steel joints. In these dissimilar joints, the main interest was to obtain a welded joint with acceptable mechanical behavior. Some papers recently focused on understanding dissimilar joints process, mainly on the metal flow and its response to corrosion. However, in Al-steel joints, the presence of steel particles in the nugget zone is routine, it alters both the mechanical and chemical behavior of welded joints. Thus, this work aims to evaluate the mechanisms that govern these particles' generation, the effect of offset on their formation, and proposing the material flow behavior, using the detached fragments as tracers. It was established that the offset controls the metal's fluidity, which allows the accumulation of steel fragments on the advanced side, and reducing its quantity, due to the decrease of irregularities in the Al-steel interface. Likewise, the metal flow was observed on the retreating side, with that mentioned in aluminum joints. In contrast, on the advanced side, there is a shear action, push down, and lateral movement towards the retreating side, driven by the high forging strength of the metal and the restriction imposed by the steel and the backing.

doi*: 10.5829/ije.2021.34.04a.29*

NOMENCLATURE

FSW	Friction stir welding	F_r	Rotational force (N)
IMC	Intermetallic compounds	P_E	Effective depth
O_r	Real offset (mm)	RS	Retreating side
O_T	Tangent offset (mm)	AS	Advancing side
P_T	Tool penetration (mm)	**Greek Symbols**	
OM	Optical microscopy	ω	Rotation speed (rpm)
SZ	Stir zone	υ	Welding speed (mm/min)
TMAZ	Thermo-mechanically affected zone	δ	slip/stick factor (mm/rev)
F_t	Travel force (N)	λ	spacing between bands (mm)

1. INTRODUCTION

Energy consumption is a notable factor in the design of new transport systems. Therefore, the reduction of vehicle weight, without compromising the integrity of the structure, is the target of many studies [1], with a focus on the automotive, [2, 3], naval [4], aeronautics [5] and aerospace [6, 7] industries. In this sense, different methodologies have emerged and continue to be evaluated. One of these is the Friction Stir Welding (FSW) process, developed by TWI in 1991 [8], which is a technique for joining and processing materials that arose from the concept of conventional friction welding. FSW uses a tool to produce heat while generating severe plastic deformation, resulting in a mechanical/metallurgical mixture of the plasticized

*Corresponding Author Email: eandres.torres@udea.edu.co (E. A. Torres)

Please cite this article as: E. A. Torres, J. Graciano-Uribe, T. F. A. Santos, Control of Steel Detachment and Metal Flow on Aluminum-steel Friction Stir Welding of Thin Joints, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1024-1034

metal [9], [10]. The tool, formed by a pin and a shoulder, has two main functions: locally heat the workpiece and stir the material to obtain the welded joint. The combined movement of rotation and translation in the tool generates relative displacement due to different speeds in both side of the tool, allowing a complex flow of metal in the plasticized zone and, consequently, consolidating a welded joint [11, 12].

Plasticized metal is limited by the tool, solid metal, and backing plate, which forces it to flow around the tool, forming the joint [13]. The flow complexity is accentuated by the tool's geometric features, such as threads and flutes, designed to drive the material and to promote adequate material mixing [14, 15]. Due to the complexity of the metal flow, different types of defects can be found, which include surface defects, voids, and lack-of-fill. The latter two are the result of low plasticity of the material and loss of filling capacity. In this sense, it was determined that the pitch (υ/ω) is an essential variable for controlling voids and kissing bonds [16].

The welding of dissimilar joints [17, 18] was approached considering three different systems: i) dissimilar joints of low melting metals, ii) the joining of hard metals, and iii) the joining of metals with different mechanical properties. The welding of systems composed of metals with very different physical properties has limitations due to the multiple challenges. The main ones are the difference in the melting temperature and mechanical strength. The FSW of dissimilar metals is different from the welding of the same metal systems by forming a more heterogeneous flow of metal [19]. Also, few references to the metal flow in the FSW of Al-steel joints make it difficult to understand and control the generation of defects.

The first experiments with welding dissimilar joints using FSW ended with developing new welding parameters, based on the flow of the plasticized material and the asymmetry of heat generation in the joint [20], which are the joint configuration and the position of the tool. The joint configuration determines how the plates should be positioned considering the advancing and retreating sides. The joint configuration for welding metals with different melting temperatures places the hardest metal on the advanced region, which is the side where the temperature generated by friction is higher [21]. The parameter that determines the position of the tool is called offset. Therefore, the offset defines the position of the tangent of the pin to the joint line, being positive when it enters steel and negative when the pin is entirely in aluminum [22, 23]. From these recommendations, Yasui [24] obtained welded joints AA6063-steel, where he related the effect of υ and ω with the plastic flow and the formation of defects, similar to what happened when welding joints of the same metal.

In addition to the complex flow of metal, the detachment of steel is usual in this type of welded joint. The majority of papers on the subject mention such fragments without going into greater detail [25]. Several works point to the formation of intermetallic compounds (IMC) around steel particles [26], while others evaluate their effect on the response to corrosion of the joint [27], being that the increment in the number of fragments significantly increases the corrosion rate [23-28]. It should be noted that different strategies have been presented to investigate the flow of material during FSW, such as insertion of markers, three-dimensional analysis using X-ray tomography or X-ray radiography in situ, use of different materials, the use of plasticine, and applied freezing after the pin breaks [29]. In one of the first works, Colligan [30] used steel markers to determine the material flow in aluminum joints. Consequently, the steel fragments detached during the FSW of Al-steel joints can be used as markers to follow the plasticized metal flow in the nugget zone.

Therefore, it is essential to understand the effect of welding parameters on the mechanisms that produce these particles and the strategies to control their formation. Furthermore, it is crucial to understand the plasticized metal flow since the detached particles' position and characteristics depend on it. This manuscript assesses the effect of offset on the formation, size, and distribution of the steel particles detached and deposited in the mixing zone. The paper also presents a proposal about the offset's influence in the aluminum plasticization process and the metal's flow in thin aluminum-steel joints welded by FSW. The authors would like to highlight the major concern about joining aluminum and steel are the variety of deleterious intermetallic compounds in the fusion zone, in case of fusion welding processes and the nugget zone when friction stir welding is used. The ability to achieve the pursued challenge, which is Al-steel joining, is connected to a better understanding and controlling the nugget zone precipitation.

2. MATERIALS AND METHODS

The materials used were aluminum alloy plates 6063-T5 and AISI SAE 1020 steel, both with dimensions of 500 × 85 × 2.0 mm. Welding was performed using a dedicated FSW machine from Transformation Technologies Incorporated (TTI), model RM-2. The machine has complete control of the welding parameters (rotation and welding speed), even with penetration of the tool controlled by position or axial load. In addition, the device has torque, liquid cooled tool holder, and wireless temperature acquisition system required for continuous real time temperature data control. A metal matrix and ceramic reinforcement tool of tungsten carbide (WC-14Co) were used, with shoulder and pin of 25 and 5.7 mm in diameter, respectively, and a pin length of 1.35 mm.

Figure 1a shows the configuration of the joint. The positioning of tool was done using two criteria: the tangent offset (O_T) and the tool real offset (O_R). The first considers the distance between the pin and the joint line's tangent, while the second holds the radius of the pin (5.7 mm) plus the displacement of the tangent (Figure 1b).

The joints were produced using ceramic backing [31]; however, to demonstrate the complete material flow, joints were made using a 5052-aluminum backing. Table 1 displays the variables and parameters used in the welding process. The joints were elaborated using the tool position control mode, with the axial force of 18 kN, to +0.5 mm offset, and 22 kN for +1.0 and +1.5 mm.

Microstructural characterization was performed using optical microscopy (OM) and scanning electron microscopy (FE-SEM). The samples were prepared using sandpaper from 100 to 1500 mesh, followed by polishing with 1.0 µm diamond paste. In order to observe the microstructure, the samples were etched with 2% nital, followed by etching with 1% hydrofluoric acid. The characterization of the steel particles was performed using the ImageJ software.

3. RESULTS

3. 1. Base Metal

Figure 2 presents the microstructure of the metals used. For both the AISI-SAE 1020 and the AA6063 the structure is composed of equiaxial grains in all directions. In the steel workpiece (Figure 2a), the structure is formed by ferrite grains with

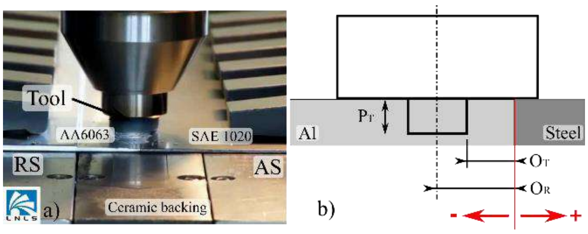


Figure 1. a) Photograph of the joint configuration, the tool and the backing. b) The positioning of the tool for determining the real (O_R) and tangent (O_T) offset. The red arrows and the positive and negative signs indicate the tool's position to the joint line. RS and AS corresponds to the retreating and advancing side, respectively

TABLE 1. Welding parameters for AA 6063-steel 1020 joints, to welding in different stages.

Test	ω (rpm)	υ (mm/min)	O_r (mm)	O_T (mm)	P_T (mm)
Ceramic backing	300	150	-2.1;-1.6;-1.1	0.5; 1.0; +1.5	1.65
Al backing	300	150	-3.1 to -0.6	-0.5 to +2.0	1.50 to 1.70

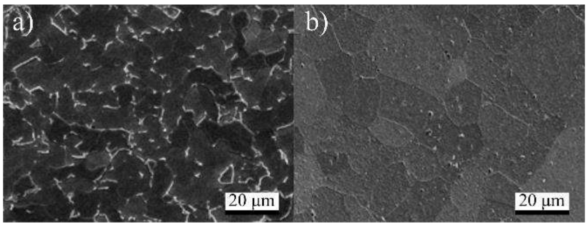


Figure 2. FE-SEM images showing the microstructure of a) AISI-SAE 1020 steel and b) AA6063-T5. Etching with nital for steel and hydrofluoric acid for aluminium

perlite (white). For the aluminum alloy (Figure 2b), the arrangement is of α_{Al} grains, with the presence of β_{AlFeSi} particles, many of them dissolved (holes) by the action of the etchant [32]. In both cases, there is no evidence of the metal rolling process.

3. 2. Macrostructure, Production, and Control of Steel Particles

Figure 3 shows the macrographs of the welded joints with O_T of +0.5 and +1.5. Two regions are easily identified on the aluminum side: nugget or stir zone (SZ), and the thermo-mechanically affected zone (TMAZ), while on the steel side, only TMAZ is evident. The welding parameters and the heat input control were essential in obtaining joints with a suitable surface appearance with no defects [33].

A similar structure, with a clear difference between the welded metals, was observed in dissimilar joints of AA7075-AA2024. The authors claim that low rotational speed negatively affects the joint, leading to absence of mixing. In contrast, satisfactory mixture is reached at high rotational speed. [34, 35]. It is not the case for aluminum-steel joints, where mixing does not occur in any condition, as highlighted by other authors [36, 37], who points out a clear limit between aluminum and steel. The only mixing occurs in the SZ, where steel fragments are observed, typical of this type of welded joint [38], [39].

The general appearance of the particles can be seen in Figure 4a. It would be plausible to believe that particles could come from the tool, but that possibility is completely ruled out. The fragments retain traces of the TMAZ, from where they were detached. For example, the cementite sheets are completely stretched by the high deformation of steel at the interface, in addition to the ferrite micro-grains formed by the dynamic recrystallization of the steel (Figure 4b). It also highlights the absence of intermetallic compounds (IMC) due to the absence of these deleterious phases in the welded joints, as it was indicated in previous work [40]. The absence of IMC contrasts those with other studies [41, 42], where these composites outline the particles, identified by Pourali et al. [43] as $FeAl_3$.

Lee et al. [44] attribute the formation of debris to the O_T's action, which leads to broken steel particles on the

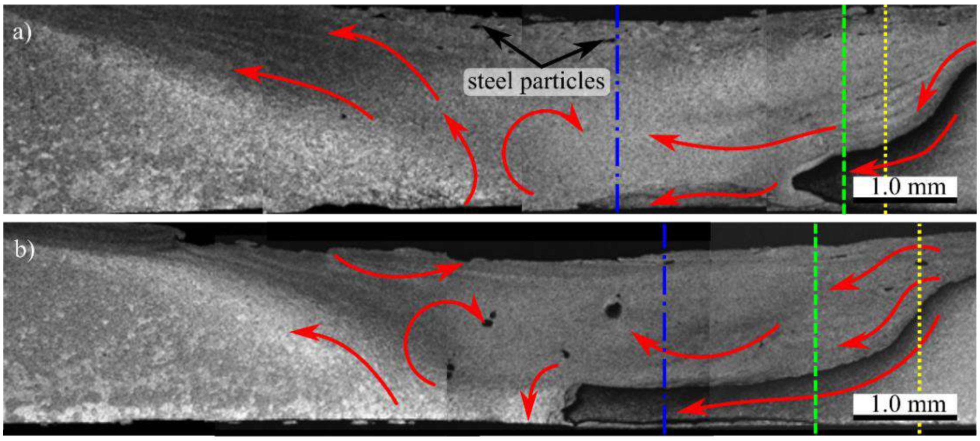


Figure 3. Macrograph of the welded joint's cross-section with O_T of a) +0.5, and b) +1.5 mm. The green, blue, and yellow lines correspond to the joint's original line, the tool axis, and the O_T, respectively. Red arrows indicate the direction of material flow

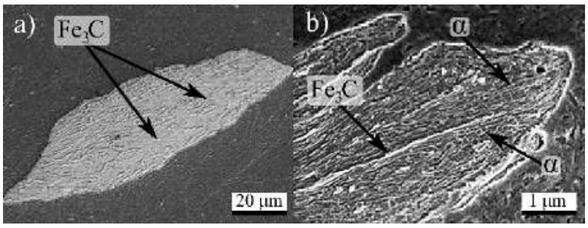


Figure 4. FE-SEM images of the particles observed in the SZ: a) big particle (180 µm) and b) detail showing the stretched lamellae of cementite and ferrite's nano-grains (α).

surface of welded joints being distributed within the SZ. Figure 5 shows images from top to bottom in joints with +0.5 and +1.5 mm offset. These images confirm the accumulation of the fragments, mainly on the surface, so it is defined that this is the most important place for its quantification.

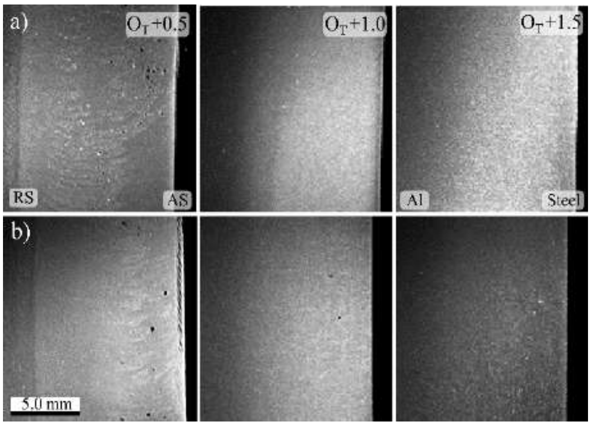


Figure 5. Distribution of steel fragments in the SZ in welded joints with offset a) +0.5 and b) +1.5 mm. The images were obtained with polished from the top at 0.5, 1.0, and 1.5 mm depth of the joint surface

The hypothesis that the quantity and size of the particles change with the O_T is being corroborated in Figure 6. Surprisingly, there is a constant reduction in the number of particles as the offset increases; the expected behavior was the opposite. Higher O_T means more interaction between tool and steel, which would easily explain the increased particle formation [45].

On the other hand, related to the steel particles size, Figure 6 does not show a discernible relationship with the offset, since it was expected that a higher O_T would generate large steel debris [46]. Figure 7 shows a more precise relation between offset and particle size by the distribution of the particle area. It confirms the reduction in the number of particles with O_T; besides, the figure registers that the particle size is less than 0.1 mm^2, for all conditions evaluated. Finally, the analysis leaves in mind that the particle size decreases with O_T, evidenced by the smaller number of particles larger than 1.0 mm^2, which confirms the reduction of both quantity and size with the offset.

Part of the formation and the detachment of the fragments are related to the generation of protuberances (Figure 8a). The tool shear stress promoted by plasticized aluminum entrance in small openings in the Al-steel

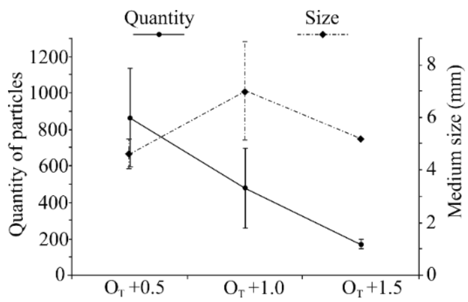


Figure 6. Effect of offset on the quantity and size of steel particles in the SZ.

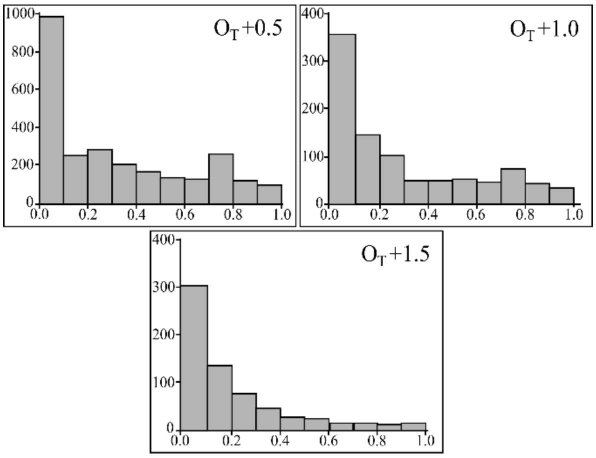


Figure 7. Measurement of the quantity and area of the steel fragments at the top of the SZ

interface (Figure 8b). Coelho et al. [47] observed the formation of similar structures that they defined as a non-smooth interface, to which the mechanical interlocking between both materials is attributed. Movahedi et al. [48] indicate that aluminum's entry favors IMC formation in a swirl-layer, formed by the mechanical mixture and the diffusion between aluminum and steel. However, this is not the case for the joints under study, as the phase in the openings corresponds completely to Al (Figure 8c).

In the flatter regions of the interface, the detachment mechanism is different, as it inVolves tearing the interface, as shown in Figure 9a, which leads to the removal of irregularities. The deformation at the interface is so high that it causes stretching and recrystallization of the ferrite grains; features are observed outside the steel fragments, as shown in Figure 9b.

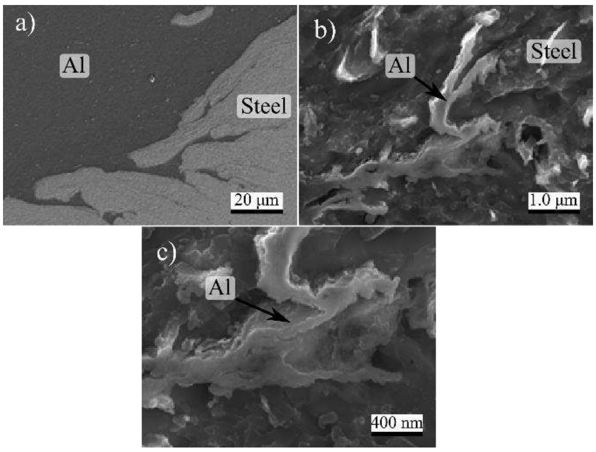


Figure 8. FE-SEM images of the final joint welded with a +1.5 mm offset. a) Formation of protuberances at the interface, b) ingress of aluminum into the steel, and c) cracking and separation of the protuberance

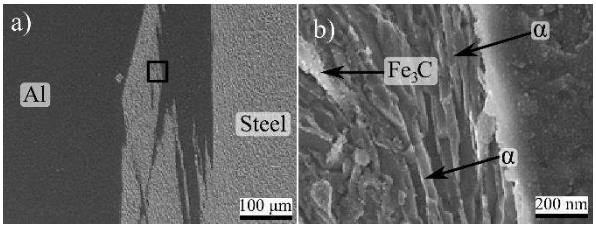


Figure 9. FE-SEM images of a) steel particle coming off the Al-steel interface and b) detail of the particle interface showing the severe deformation of the ferrite grains

3. 3. Metal Flow Figure 3 exposes the flow of material in the thin sheet Al-steel joint. Metal flow was described, considering morphological features observed in the macrographs such as banding, steel's profile at the interface, the shape of the SZ, and the steel fragments location. Kimapong and Watanabe [49] indicate that the particles aligned with the aluminum flow, as observed in Figure 10, follow the plasticized metal movement.

Banding is one of the most prominent peculiarities in the SZ of metals processed by FSW. In this process, the plasticized metal displacement occurs both in a laminar and vertical way. As the tool advances, plasticized metal is added layer-by-layer to the joint's back, which generates the banding, better known as the onion ring structure [50]. For the joints in question, this is more noticeable in regions with a considerable accumulation of

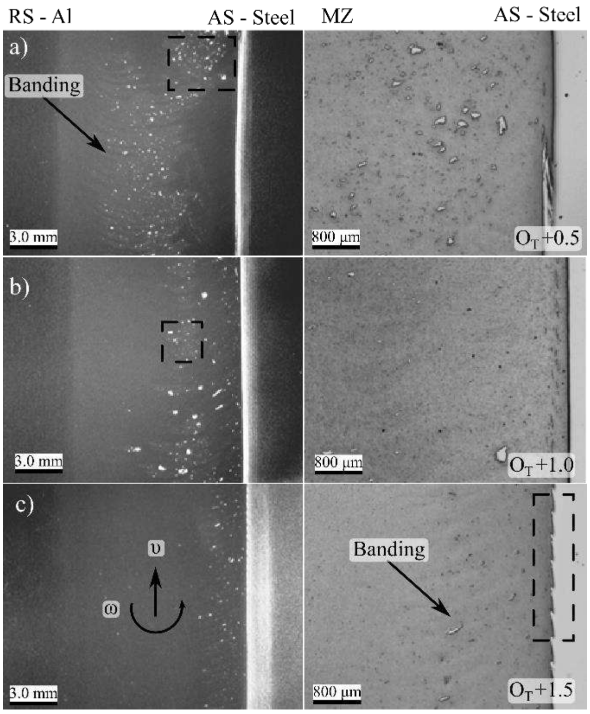


Figure 10. Macrographs at the top of the joints with an offset of a) +0.5; b) +1.0 and c) +1.5 mm.

steel fragments. Fonda and Bingert [51] established that these bands correspond to structural variations such as grain size, particle distribution, or texture, while multiple authors [52-55] report that the spacing between bands (λ) corresponds to the advance per reVolution (υ/ω).

Another critical point is the position of the particles. As presented in Figure 10, as the O_T increases, the steel fragments accumulate on the advanced side. This phenomenon can be justified considering the scheme of Figure 11a. Chen et al. [56] explain that the so-called shear zone occurs at the front of the pin, gradually growing as it moves towards the RS. The material transferred from the AS is moved towards the pin's back, generating a layer, which forms the banding. The highly deformed material flows around the pin forming the swirl zones. It is responsible for forming of defects such as voids when the plasticized metal's speed is lower to reach the metal at the rear of the advanced side [57]. Kumar and Kailas [58] point out that void defects are eliminated as welding forces increase, as the metal's extrusion force increases. However, other authors indicate that voids presence is the result of the lack of adherence between the plasticized metal and the pin [59, 60]. Kumar and Kailas [61] showed that for small O_T, much of the entrained material is deposited at the rear of the advanced side, which explains the position of the steel particles.

In a joint of the same material, the interaction between the tool and the pieces generates different forces represented in Figure 11b. Two types of forces can be highlighted: the normal force (F_n), generated by the forward movement of the tool, and the shear force (F_s) produced by the friction between the pin and the metal, where their direction and magnitude vary with the tool's position. Coelho et al. [62] agree with this approach but explain that they are two fundamental forces: the advanced force (F_{travel}) and the rotational force ($F_{rotation}$). Such forces are added or subtracted at some point, generating F_n and F_s; thus, the forces magnitude depends on the welding parameters and the welded metal's mechanical and physical properties.

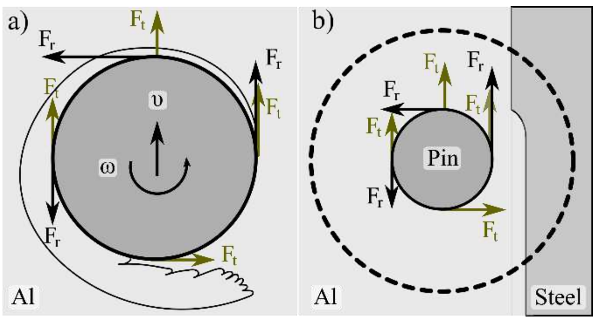


Figure 11. a) Diagram of the material flow in the shear zone around the pin (adapted from Chen et al. [56]). b) Diagram of the forces acting on the tool (adapted from Coelho et al. [62]).

Therefore, the shear force on the feed side ($F_{s(adv)}$) is the sum of F_{travel} and $F_{rotation}$, which generates a region of high forging pressure, accentuated by the constriction of the shoulder and the steel. This force is high enough to promote the steel's significant deformation and generate its recrystallization at the interface (Figure 9). When the force is very high, this leads to the stirring pin into the steel surface, producing the so-called fin-like shaper [63] observed in the rectangle in Figure 10c. This pressure on the advanced side pushes the steel down. However, the ceramic backing reacts against this movement, forcing the metal to move horizontally, below the pin, towards the aluminum side (Figure 3).

The deformation of the steel is different in the case of an aluminum backing. Figure 12 presents the welded joints results with the backing of AA5052, which shows the metal's full flow, without the restriction imposed by a higher hardness backing. This figure reveals how the steel moves and goes beyond the pin from the advanced to the retreating side. The effective depth (P_E) increases with the offset in joints welded with the same tool penetration (P_T). The flow of metal pushes the steel to the bottom from the AS, subsequently carried upwards when it reaches the RS. In this case, the upward movement of material is caused by the advanced side's metal flow.

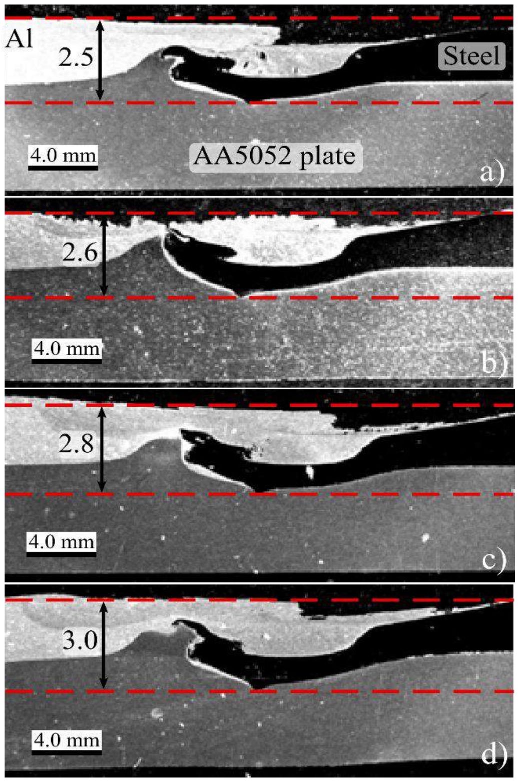


Figure 12. Pictures of FSW's aluminum-steel joint, using P_T 1.8 mm, O_T -0.3, +0.5, +1.3, and +2.0 mm, and AA5052 backing plate. The distance between the dashed lines (red) corresponds to the weld's P_E (in mm)

The upward movement of the steel on the RS (Figure 12) is the same that pushes the aluminum upward, on the same side at Al-steel joints (Figure 3). On this side, F_s is lower, since the advanced movement of the tool is opposite to the rotation, generating a low-pressure area, which favors the flow from the bottom to the surface.

4. DISCUSSION

The welding zone's morphological features are inherent to metal's flow, both in solid and plasticized state. One of the models in FSW considers the material flow as the sum of three combined movements: 1) a cylindrical flow around the pin, 2) a homogeneous flow parallels to the welding direction, and 3) an upward vortex-shaped flow around the pin [64]. Fonda et al. [65] explain that the flow around the pin is produced by the tool's rotation, which results in the shear deformation that originates from the shear zone. Likewise, they point out that the downward and upward movement of metal (vortex) is due to the threads and the tool's translation, generating the "onion ring". The spacing between bands corresponds to the tool advance per reVolution. Gerlich et al. [66] e Avettand-Fènoël et al. [67] also consider that the SZ was made up of only two flows: 1) the flow generated by the pin and 2) that produced by the axial force and rotation of the shoulder [50]. However, these theories are devised for the flow of metal in joints of the same metal. A few works make proposals for dissimilar joints between a soft metal and another of high resistance. Nevertheless, these elements must be considered to explain the observed behavior about the shape, distribution of steel particles in the SZ, and propose a model for FSW in thin Al-steel joints.

During the initial contact of the tool with the joint, the steel particles detachment occurs due to the shoulder's erosive action. The high pressure and the displacement of the plasticized aluminum cause the deformation of the wedge-shaped steel (Figure 3b). Texier et al. [68] point out that the intensity or speed at which extrusion occurs depends on the material's relative position to the pin. Doude et al. [69] studied the combined effect of shoulder and pin movements, added to the restriction imposed by backing, on the generation of symmetric vortex flow in the SZ (Figure 13a). However, in Al-steel joints, the symmetry is broken. The steel acts as another barrier to metal flow, altering the distribution of the plasticized material, allowing only its downward movement into the AS. In turn, this pushes the steel towards the bottom, to the pin's tip, and forces the steel to move horizontally in the direction of the RS (Figure 13b). The degree of deformation of the steel is subject to offset since this factor controls magnitude of the forging force. Wan and Huang [70] came to a similar result, where the forging force increased by the tool plunge.

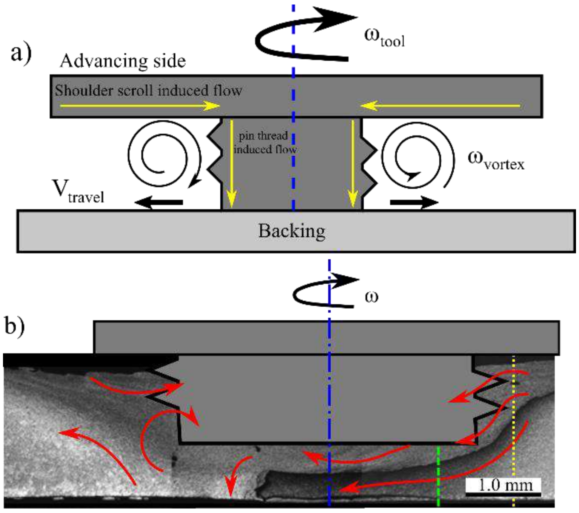


Figure 13. a) Scheme of the vortex generation in FSW in an aluminum joint and b) plot of the vortex and displacement of the metal in the Al-steel joint.

On the other hand, the steel particles detachment was produced by the interaction between the shoulder and the pushing force of plasticized metal removing irregularities from the steel surface, leaving it smoother. The shear movement generated by the pin promotes the emergence of irregularities and openings in the interface. However, as the offset increases, the forging pressure between the pin-shoulder and the steel is high enough to flatten the surface, reducing the number of irregularities and the detached particles mass. For this, the forging pressure close to the possible openings in the Al-steel interface, preventing the entrance of the plasticized metal. Smaller fragments come off since the irregularities are also smaller.

Most of the particles were formed by the contact between the shoulder and the joint surface. Few particles are dragged downwards, as shown by the series of images in Figure 5. Particles displaced by shoulder remain in their influences the area, while those caused inside the SZ are displaced to the surface by vortex movement. From this, it is essential to establish a link between the offset and the particles' position.

Liechty and Webb [71] determined flow lines in FSW, the result of which is superimposed on the macrographs of the deposits, as presented in Figure 14. A fraction of the pin interacts with the steel with a small offset, but such coincides with the flow that rotates more than 180° around the pin, allowing the fragments to pass from the AS to the RS and continue to the rear of the pin (Figure 14a). As O_T increases, many of the AS flow lines move parallel to the weld axis so that the detached particles are not trapped by the pin flow, crossing by the same AS (Figure 14b-c).

In another proposal for metal flow, Zeng et al. [72] relate the displacement of the plasticized metal with the joint's temperature since it defines the degree of fluidity of the metal. Figures 14a-c shows Zeng's results superimposed with the particle distribution to +0.5, and +1.5 mm offset. For this type of joints, it was determined that both heat input and maximum temperature increase with the offset [73, 74], and in that sense, there is an agreement with Zang's proposal and the observed results.

In FSW, heat is the main product of friction between the shoulder and the joint [75], the shear deformation produced by the pin [76]. Therefore, the heat generation depends on contact conditions of the elements since the metal can stick or slide on the tool [77]. Idagawa et al. [78] relate the slip/stick conditions with the offset and the

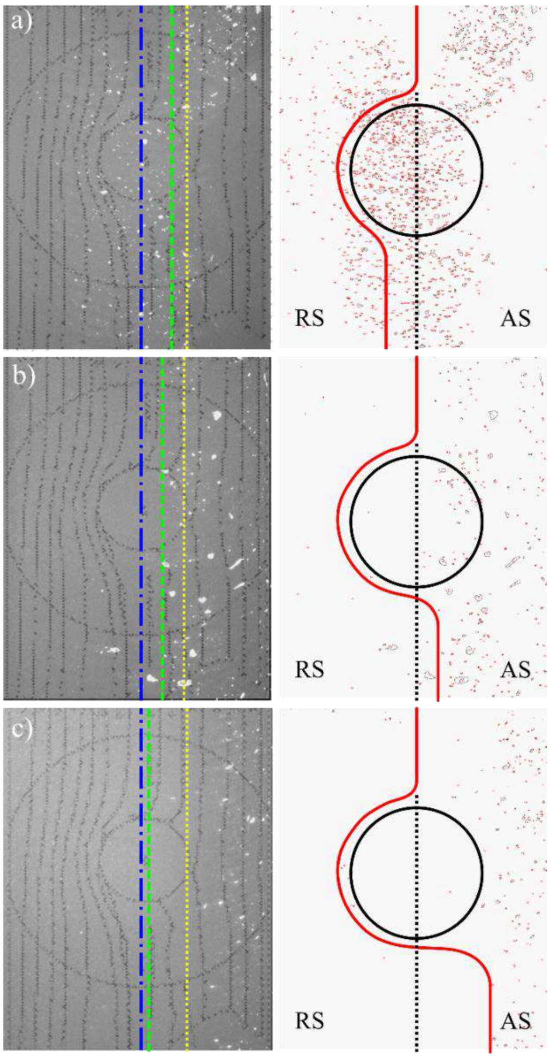


Figure 14. Overlapping between steel particle distribution and metal flow model proposed by Liechty and Webb [68], for offset: a) +0.5, b) +1.0, and c) +1.5 mm. Projection connecting the distribution of steel particles with the metal flow model proposed by Zeng et al. [69], for offset: a) +0.5, b) +1.0, and c) +1.5 mm. The green, blue, and yellow lines correspond to the original joint line, the tool axis, and the O_T, respectively

heat generation in these joints, through the slip/stick factor (δ) for each material (δ_{Al} and δ_{Steel}). If $\delta = 1$, there is virtually no adherence between metal and tool, so the heat is generated mainly by friction.

On the other hand, if $\delta = 0$, the heat is produced almost entirely by plastic deformation [79], [Souza et al. DOI: 10.3217/978-3-85125-615-4-33]]. Idagawa was able to establish that, on the aluminum side, the predominant mechanism for heat generation is plastic deformation, which implies that plasticized metal adheres to the tool, while for steel, the friction responds for 85% of the heat produced, consequently prevail slip on the steel-tool interaction. Even more important was to establish that δAl increases significantly with the offset, going from 0.02 to 0.40, meaning that adhesion is lost between the elements because the plasticized metal reaches significant fluidity. Therefore, this sings that the detached particles' location is related to the metal's fluidity, which depends on the joint's temperature, which in the case in question is subject to offset.

5. CONCLUSIONS

This work evaluated the generation and distribution of steel fragments and the metal flow in thin sheets of aluminum-steel welded joints. From the results and their analysis, the following points are concluded.

The detachment of steel particles was produced by two mechanisms: generation of protuberances at the interface by introducing plasticized aluminum and the tearing of the surface by the shear stress of the aluminum flow.

Detached particles correspond only to steel since they have cementite sheets and ultra-fine grains of ferrite, generated by the steel's high deformation and dynamic recrystallization.

The quantity of fragments decreases with the offset due to the reduction in protuberances formation by increasing the forging force, which inhibits plasticized aluminum entrance at the interface.

Forging force at AS increases with the offset, which implies that the steel's surface moves down and below the pin, where the restriction of the backing forces its horizontal displacement towards the RS.

The particles position is defined by the offset, which controls the plasticized fluidity of metals by determining the temperature in the joint. Large offset generates higher temperature and metal fluidity, allowing that particles to be dragged to the pin's back, closer to AS's interface. Meanwhile, a small offset decreases the temperature and fluidity of aluminums, promoting that the fragments are led to the pin's back, closer to the centerline.

The containment between the shoulder, pin, and backing, on the retreating side, promote the material to flow in a vortex shape. In contrast, on the advancing side, the flow's symmetry is broken by the restriction imposed by the steel, which promotes shear and downward movement of the plasticized metal.

6. REFERENCES

1 S. Sheikhi and J. F. Dos Santos, "Effect of process parameter on mechanical properties of friction stir welded tailored blanks from aluminium alloy 6181-T4," ***Science and Technology of Welding and Joining***, Vol. 12, No. 4, 370-375, (2007), doi: 10.1179/174329307X173698.

2 K. Colligan, P. Konkol, J. Fisher, and J. Pickens, "Improved tools and process parameters were used to fabricate structures of 2519 aluminum armor for the U.S. Marine Corps' Advanced Amphibious Assault Vehicle," ***Welding Journal***, Vol. 82, No. 3, (2003).

3 M. Posada, J. P. Nguyen, D. R. Forrest, J. J. DeLoach, and R. Denale, "Friction stir welding advances joining technology," ***Amptiac Q.***, Vol. 7, No. 3, 13-20, (2003).

4 K. Colligan, "Friction stir welding for ship construction," ***Concurrent Technologies Corporation, Harrisburg, PA***. 2004, [Online]. Available: http://147.160.99.83/useruploads/file/publications/FSWShipConstruction.pdf.

5 S. W. Williams, "Welding of airframes using friction stir," Air Sp. Eur., Vol. 3, No. 3-4, 64-66, (2001), doi: 10.1016/s1290-0958(01)90059-0.

6 M. J. Brooker, A. J. M. Van Deudekom, S. W. Kallee, and P. D. Sketchley, "Applying friction stir welding to the Ariane 5 main motor thrust frame," European Space Agency, (Special Publication) ESA SP, No. 468. 507-511, (2001).

7 J. Ding, R. Carter, K. LA WLESS, A. Nunes, and C. Russell, "Friction stir welding flies high at NASA," ***Welding Journal***, Vol. 85, No. 3, 54-59, (2006).

8 W. M. Thomas, E. D. Nicholas, J. C. Needhan, M. G. Murch, P. Temple-Smith, and C. J. Dawes, "International patent application PCT/GB92/02203 and GB patent application 9125978.8," UK Patent Office, London. 1991.

9 R. S. Mishra and Z. Y. Ma, "Friction stir welding and processing," ***Materials Science and Engineering R: Reports***, Vol. 50, No. 1-2. Elsevier, 1-78, (2005), doi: 10.1016/j.mser.2005.07.001.

10 I. T. C. ; J. Langari, F. Kolahan, and K. Aliakbari, "Effect of Tool Speed on Axial Force, Mechanical Properties and Weld Morphology of Friction Stir Welded Joints of A7075-T651," ***International Journal of Engineering, Transactions C: Aspects***, Vol. 29, No. 3, 403-410, (2016), doi: 10.5829/idosi.ije.2016.29.03c.15.

11 J. Q. Su, T. W. Nelson, R. Mishra, and M. Mahoney, "Microstructural investigation of friction stir welded 7050-T651 aluminium," ***Acta Materialia***, Vol. 51, No. 3, 713-729, (2003), doi: 10.1016/S1359-6454(02)00449-4.

12 R. Singh, S. A. Rizvi, and S. P. Tewari, "Effect of Friction Stir Welding on the Tensile Properties of AA6063 under Different Conditions," ***International Journal of Engineering, Transactions A: Basics***, Vol. 30, No. 4, 597-603, (2017).

13 R. M. Leal, C. Leitão, A. Loureiro, D. M. Rodrigues, and P. Vilaça, "Material flow in heterogeneous friction stir welding of thin aluminium sheets: Effect of shoulder geometry," ***Materials Science and Engineering A***, Vol. 498, No. 1-2, 384-391, 2008, doi: 10.1016/j.msea.2008.08.018.

14 P. B. Prangnell and C. P. Heason, "Grain structure formation during friction stir welding observed by the 'stop action technique,'" ***Acta Materialia***, Vol. 53, No. 11, 3179-3192, (2005), doi: 10.1016/j.actamat.2005.03.044.

15 I. Transactions B ; J. Fabregas, A. Villegas, and J. Martínez Guarín, "A Coupled Rigid-viscoplastic Numerical Modeling for Evaluating Effects of Shoulder Geometry on Friction Stir-welded Aluminum Alloys," ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 2, 313-321, (2019), doi: 10.5829/ije.2019.32.02b.17.

16 S. Xu and X. Deng, "A study of texture patterns in friction stir welds," ***Acta Materialia***, Vol. 56, No. 6, 1326-1341, 2008, doi: 10.1016/j.actamat.2007.11.016.

17 O. D. Hincapié, J. A. Salazar, J. J. Restrepo, E. A. Torres, and J. Graciano-Uribe, "Control of Formation of Intermetallic Compound in Dissimilar Joints Aluminum-steel," ***International Journal of Engineering, Transactions A: Basics***, Vol. 32, No. 1, 127-136, (2019), Accessed: Jul. 13, 2019. [Online].

18 N. Ethiraj, T. Sivabalan, B. Sivakumar, S. Vignesh Amar, N. Vengadeswaran, and K. Vetrivel, "Effect of tool rotational speed on the tensile and microstructural properties of friction stir welded different grades of stainless steel joints," ." ***International Journal of Engineering, Transactions A: Basics***, Vol. 33, No. 1, 141-147, (2020), doi: 10.5829/ije.2020.33.01a.16.

19 L. E. Murr, "A review of FSW research on dissimilar metal and alloy systems," ***Journal of Materials Engineering and Performance***, Vol. 19, No. 8, 1071-1089, (2010), doi: 10.1007/s11665-010-9598-0.

20 T. DebRoy and H. K. D. H. Bhadeshia, "Friction stir welding of dissimilar alloys - A perspective," ***Science and Technology of Welding and Joining***, Vol. 15, No. 4, 266-270, (2010), doi: 10.1179/174329310X12726496072400.

21 V. Firouzdor and S. Kou, "Al-to-Mg friction stir welding: Effect of material position, travel speed, and rotation speed," ***Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 41, No. 11, 2914-2935, (2010), doi: 10.1007/s11661-010-0340-1.

22 T. Kakiuchi, Y. Uematsu, and K. Suzuki, "Evaluation of fatigue crack propagation in dissimilar Al/steel friction stir welds," **in *Procedia Structural Integrity***, (2016), Vol. 2, 1007-1014, doi: 10.1016/j.prostr.2016.06.129.

23 S. Y. Anaman, H. H. Cho, H. Das, J. S. Lee, and S. T. Hong, "Microstructure and mechanical/electrochemical properties of friction stir butt welded joint of dissimilar aluminum and steel alloys," ***Materials Characterization***, Vol. 154, 67-79, (2019), doi: 10.1016/j.matchar.2019.05.041.

24 T. Yasui, T. Ishii, Y. Shimoda, M. Tsubaki, M. Fukumoto, and T. Shinoda, "Friction stir welding between aluminum and steel with high welding speed," ***Production.*** 2004.

25 Q. Zheng, X. Feng, Y. Shen, G. Huang, and P. Zhao, "Dissimilar friction stir welding of 6061 Al to 316 stainless steel using Zn as a filler metal," ***Journal of Alloys and Compounds***, Vol. 686, 693-701, (2016), doi: 10.1016/j.jallcom.2016.06.092.

26 T. Wang, H. Sidhar, R. S. Mishra, Y. Hovanski, P. Upadhyay, and B. Carlson, "Evaluation of intermetallic compound layer at aluminum/steel interface joined by friction stir scribe technology," ***Materials and Design***, Vol. 174, 107795, (2019), doi: 10.1016/j.matdes.2019.107795.

27 B. Seo, K. H. Song, and K. Park, "Corrosion Properties of Dissimilar Friction Stir Welded 6061 Aluminum and HT590 Steel," ***Metals and Materials International***, Vol. 24, No. 6, 1232-1240, (2018), doi: 10.1007/s12540-018-0135-2.

28 M. Thomä, G. Wagner, B. Straß, B. Wolter, S. Benfer, and W. Fürbeth, "Ultrasound enhanced friction stir welding of aluminum and steel: Process and properties of EN AW 6061/DC04-Joints," ***Journal of Materials Science and Technology***, Vol. 34, No. 1, 163-172, (2018), doi: 10.1016/j.jmst.2017.10.022.

29 Y. Huang *et al.*, "Material-flow behavior during friction-stir welding of 6082-T6 aluminum alloy," ***The International Journal of Advanced Manufacturing Technology***, Vol. 87, No. 1-4,

1115-1123, (2016), doi: 10.1007/s00170-016-8603-7.

30 K. Colligan, "Material flow behavior during friction stir welding of aluminum," ***Weld. Journal-New York-***, Vol. 78, 229-s, 1999.

31 T. F. A. Santos, E. A. Torres, T. F. C. Hermengildo, and A. J. Ramirez, "Development of ceramic backing for friction stir welding and processing," ***Welding International***, Vol. 30, No. 5, 338-347, (2016), doi: 10.1080/09507116.2015.1096498.

32 E. Torres and A. Ramirez, "União de juntas dissimilares alumínio-aço de chapas finas pelo processo de soldagem por atrito com pino não consumível (sapnc)," Soldag. e Insp., Vol. 16, No. 3, 265-273, (2011), doi: 10.1590/S0104-92242011000300008.

33 E. Torres and A. Ramirez, "Efeito dos parâmetros de processo na obtenção e na microestrutura de juntas alumínio-aço realizadas mediante Soldagem POR Atrito COM Pino nãO Consumível (SAPNC)," Soldag. e Insp., Vol. 18, No. 3, 245-256, (2013), doi: 10.1590/S0104-92242013000300007.

34 P. Alvarez, G. Janeiro, A. A. M. Da Silva, E. Aldanondo, and A. Echeverría, "Material flow and mixing patterns during dissimilar FSW," in ***Science and Technology of Welding and Joining***, (2010), Vol. 15, No. 8, 648-653, doi: 10.1179/136217110X12785889549543.

35 A. A. M. da Silva, E. Arruti, G. Janeiro, E. Aldanondo, P. Alvarez, and A. Echeverria, "Material flow and mechanical behaviour of dissimilar AA2024-T3 and AA7075-T6 aluminium alloys friction stir welds," ***Materials and Design***, Vol. 32, No. 4, 2021-2027, (2011), doi: 10.1016/j.matdes.2010.11.059.

36 W. H. Jiang and R. Kovacevic, "Feasibility study of friction stir welding of 6061-T6 aluminium alloy with AISI 1018 steel," *Proc. **Inst. Mech. Eng. Part B J. Eng. Manuf***., Vol. 218, No. 10, 1323-1331, 2004, doi: 10.1243/0954405042323612.

37 H. Uzun, C. Dalle Donne, A. Argagnotto, T. Ghidini, and C. Gambaro, "Friction stir welding of dissimilar Al 6013-T4 To X5CrNi18-10 stainless steel," ***Materials and Design***, Vol. 26, No. 1, 41-46, (2005), doi: 10.1016/j.matdes.2004.04.002.

38 S. Amini and M. R. Amiri, "Study of ultrasonic vibrations' effect on friction stir welding," ***The International Journal of Advanced Manufacturing Technology***, Vol. 73, No. 1-4, 127-135, 2014, doi: 10.1007/s00170-014-5806-7.

39 R. Rafiei, A. Ostovari Moghaddam, M. R. Hatami, F. Khodabakhshi, A. Abdolahzadeh, and A. Shokuhfar, "Microstructural characteristics and mechanical properties of the dissimilar friction-stir butt welds between an Al-Mg alloy and A316L stainless steel," ***The International Journal of Advanced Manufacturing Technology***, Vol. 90, No. 9-12, 2785-2801, (2017), doi: 10.1007/s00170-016-9597-x.

40 E. A. Torres López and A. J. Ramirez, "Inhibition of the formation of intermetallic compounds in aluminum-steel welded joints by friction stir welding," ***Revista de Metalurgia,*** Vol. 51, No. 4, 1-10, 2015, doi: 10.3989/revmetalm.053.

41 T. Chen, "Process parameters study on FSW joint of dissimilar metals for aluminum-steel," ***Journal* of Materials *Science***, Vol. 44, No. 10, 2573-2580, (2009), doi: 10.1007/s10853-009-3336-8.

42 X. Liu, S. Lan, and J. Ni, "Analysis of process parameters effects on friction stir welding of dissimilar aluminum alloy to advanced high strength steel," ***Materials and Design***, Vol. 59, 50-62, 2014, doi: 10.1016/j.matdes.2014.02.003.

43 M. Pourali, A. Abdollah-zadeh, T. Saeid, and F. Kargar, "Influence of welding parameters on intermetallic compounds formation in dissimilar steel/aluminum friction stir welds," ***J. Alloys Compd.***, Vol. 715, 1-8, (2017), doi: 10.1016/j.jallcom.2017.04.272.

44 W. B. Lee, M. Schmuecker, U. A. Mercardo, G. Biallas, and S. B. Jung, "Interfacial reaction in steel-aluminum joints made by friction stir welding," ***Scripta Materialia*** , Vol. 55, No. 4, 355-358, (2006), doi: 10.1016/j.scriptamat.2006.04.028.

45 A. A. Fallahi, A. Shokuhfar, A. Ostovari Moghaddam, and A. Abdolahzadeh, "Analysis of SiC nano-powder effects on friction stir welding of dissimilar Al-Mg alloy to A316L stainless steel," ***Journal of Manufacturing Processes***, Vol. 30, 418-430, (2017), doi: 10.1016/j.jmapro.2017.09.027.

46 X. Liu, S. Lan, and J. Ni, "Experimental investigation on joining dissimilar aluminum alloy 6061 to TRIP 780/800 steel through friction stir welding," ***Journal of Engineering Materials and Technology, Transactions of the ASME***, Vol. 137, No. 4, 2015, doi: 10.1115/1.4030480.

47 R. S. Coelho, A. Kostka, J. F. dos Santos, and A. Kaysser-Pyzalla, "Friction-stir dissimilar welding of aluminium alloy to high strength steels: Mechanical properties and their relation to microstructure," ***Materials Science and Engineering A*** , Vol. 556, 175-183, (2012), doi: 10.1016/j.msea.2012.06.076.

[48 M. Movahedi, A. H. Kokabi, S. M. Seyed Reihani, and H. Najafi, "Effect of tool travel and rotation speeds on weld zone defects and joint strength of aluminium steel lap joints made by friction stir welding," ***Science and Technology of Welding and Joining***, Vol. 17, No. 2, 162-167, (2012), doi: 10.1179/1362171811Y.0000000092.

49 K. Kimapong and T. Watanabe, "Friction stir welding of aluminum alloy to steel," ***Welding Journal***, Vol. 83, No. 10, p. 277, 2004.

50 F. Gratecap, M. Girard, S. Marya, and G. Racineux, "Exploring material flow in friction stir welding: Tool eccentricity and formation of banded structures," ***International Journal of Material Forming***, Vol. 5, No. 2, 99-107, (2012), doi: 10.1007/s12289-010-1008-5.

51 R. W. Fonda and J. F. Bingert, "Texture variations in an aluminum friction stir weld," ***Scripta Materialia*** , Vol. 57, No. 11, 1052-1055, (2007), doi: 10.1016/j.scriptamat.2007.06.068.

52 S. Muthukumaran and S. K. Mukherjee, "Multi-layered metal flow and formation of onion rings in friction stir welds," ***The International Journal of Advanced Manufacturing Technology***, Vol. 38, No. 1-2, 68-73, (2008).

53 Z. W. Chen and S. Cui, "On the forming mechanism of banded structures in aluminium alloy friction stir welds," ***Scripta Materialia ,*** Vol. 58, No. 5, 417-420, Mar. (2008), doi: 10.1016/j.scriptamat.2007.10.026.

54 J. Teimournezhad and A. Masoumi, "Experimental investigation of onion ring structure formation in friction stir butt welds of copper plates produced by non-threaded tool pin," ***Science and Technology of Welding and Joining***, Vol. 15, No. 2, 166-170, 2010, doi: 10.1179/136217109X12577814486610.

55 J. Schneider, S. Brooke, and A. C. Nunes, "Material Flow Modification in a FSW Through Introduction of Flats," ***Metall. Mater. Trans. B Process Metall. Mater. Process.*** *Sci*., Vol. 47, No. 1, 720-730, (2016), doi: 10.1007/s11663-015-0523-7.

56 Z. W. Chen, T. Pasang, and Y. Qi, "Shear flow and formation of Nugget zone during friction stir welding of aluminium alloy 5083-O," ***Materials Science and Engineering A ,*** Vol. 474, No. 1-2, 312-316, (2008), doi: 10.1016/j.msea.2007.05.074.

57 A. Tongne, C. Desrayaud, M. Jahazi, and E. Feulvarch, "On material flow in Friction Stir Welded Al alloys," ***Journal of Materials Processing Technology***, Vol. 239, 284-296, (2017), doi: 10.1016/j.jmatprotec.2016.08.030.

58 K. Kumar and S. V. Kailas, "The role of friction stir welding tool on material flow and weld formation," ***Materials Science and Engineering A*** , Vol. 485, No. 1-2, 367-374, (2008), doi: 10.1016/j.msea.2007.08.013.

59 R. Zettler *et al.*, "Effect of tool geometry and process parameters on material flow in FSW of AN AA 2024-T351 alloy," in ***Welding in the World,*** (2005), Vol. 49, No. 3-4, 41-46, doi: 10.1007/BF03266474.

60 Y. H. Zhao, S. B. Lin, F. X. Qu, and L. Wu, "Influence of pin geometry on material flow in friction stir welding process," ***Materials Science and Technology*** , Vol. 22, No. 1, 45-50, (2006), doi: 10.1179/174328406X78424.

61 K. Kumar and S. V. Kailas, "Positional dependence of material flow in friction stir welding: Analysis of joint line remnant and its relevance to dissimilar metal welding," ***Science and Technology of Welding and Joining***, Vol. 15, No. 4, 305-311, (2010), doi: 10.1179/136217109X12568132624280.

62 R. S. Coelho, A. Kostka, J. dos Santos, and A. R. Pyzalla, "EBSD Technique Visualization of Material Flow in Aluminum to Steel Friction-stir Dissimilar Welding," ***Advanced Engineering Materials***, Vol. 10, No. 12, 1127-1133, (2008), doi: 10.1002/adem.200800227.

[63 K. Kimapong and T. Watanabe, "Lap Joint of A5083 Aluminum Alloy and SS400 Steel by Friction Stir Welding," ***Materials Transactions*** Vol. 46, No. 4, 835-841, (2005), doi: 10.2320/matertrans.46.835.

64 J. A. Schneider and A. C. Nunes, "Characterization of plastic flow and resulting microtextures in a friction stir weld," ***Metallurgical and Materials Transactions B Process Metallurgy and Materials Science***, Vol. 35, No. 4, 777-783, 2004, doi: 10.1007/s11663-004-0018-4.

65 R. Fonda, A. Reynolds, C. R. Feng, K. Knipling, and D. Rowenhorst, "Material flow in friction stir welds," ***Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 44, No. 1, 337-344, (2013), doi: 10.1007/s11661-012-1460-6.

66 A. Gerlich, P. Su, M. Yamamoto, and T. H. North, "Material flow and intermixing during dissimilar friction stir welding," ***Science and Technology of Welding and Joining***, Vol. 13, No. 3, 254-264, (2008), doi: 10.1179/174329308X283910.

67 M. N. Avettand-Fenoel, R. Taillard, J. Laye, and T. O. Vre, "Experimental investigation of three-dimensional (3-D) material flow pattern in thick dissimilar 2050 friction-stir welds," ***Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 45, No. 2, 563-578, 2014, doi: 10.1007/s11661-013-2015-1.

68 D. Texier, Y. Zedan, T. Amoros, E. Feulvarch, J. C. Stinville, and P. Bocher, "Near-surface mechanical heterogeneities in a dissimilar aluminum alloys friction stir welded joint," ***Materials and Design***, Vol. 108, 217-229, (2016), doi: 10.1016/j.matdes.2016.06.091.

69 H. R. Doude, J. A. Schneider, and A. C. Nunes, "Influence of the tool shoulder contact conditions on the material flow during friction stir welding," ***M Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 45, No. 10, 4411-4422, 2014, doi: 10.1007/s11661-014-2384-0.

70 L. Wan and Y. Huang, "Friction stir welding of dissimilar aluminum alloys and steels: a review," ***The International Journal of Advanced Manufacturing Technology***, Vol. 99, No. 5-8, 1781-1811, (2018), doi: 10.1007/s00170-018-2601-x.

71 B. C. Liechty and B. W. Webb, "Flow field characterization of friction stir processing using a particle-grid method," ***Journal of Materials Processing Technology***, Vol. 208, No. 1-3, 431-443, (2008), doi: 10.1016/j.jmatprotec.2008.01.008.

72 X. H. Zeng, P. Xue, D. Wang, D. R. Ni, B. L. Xiao, and Z. Y. Ma, "Effect of Processing Parameters on Plastic Flow and Defect Formation in Friction-Stir-Welded Aluminum Alloy," ***Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 49, No. 7, 2673-2683, (2018), doi: 10.1007/s11661-018-4615-2.

73 X. Liu, G. Chen, J. Ni, and Z. Feng, "Computational Fluid Dynamics Modeling on Steady-State Friction Stir Welding of Aluminum Alloy 6061 to TRIP Steel," ***Journal of Manufacturing Science and Engineering, Transactions of the ASME***, Vol. 139, No. 5, (2017), doi: 10.1115/1.4034895.

74 K. Chen, X. Liu, and J. Ni, "A review of friction stir-based processes for joining dissimilar materials," ***The International Journal of Advanced Manufacturing Technology***, Vol. 104, No. 5-8, 1709-1731, (2019), doi: 10.1007/s00170-019-03975-w.

75 Frigaard, Grong, and O. T. Midling, "A process model for friction stir welding of age hardening aluminum alloys," ***Metallurgical and Materials Transactions A Physical Metallurgy and Materials Science***, Vol. 32, No. 5, 1189-1200, (2001), doi: 10.1007/s11661-001-0128-4.

76 M. Z. H. Khandkar, J. A. Khan, and A. P. Reynolds, "Prediction of temperature distribution and thermal history during friction stir welding: input torque based model," ***Science and Technology of Welding and Joining***, Vol. 8, No. 3, 165-174, (2003).

77 J. Schneider, R. Beshears, and A. C. Nunes, "Interfacial sticking and slipping in the friction stir welding process," ***Materials Science and Engineering A*** , Vol. 435-436, 297-304, (2006), doi: 10.1016/j.msea.2006.07.082.

78 H. S. Idagawa, E. A. Torres, and A. J. Ramirez, "CFD modeling of dissimilar aluminum-steel friction stir welds," (2013).

79 R. Nandan, T. DebRoy, and H. K. D. H. Bhadeshia, "Recent advances in friction-stir welding - Process, weldment structure and properties," ***Progress in Materials Science***, Vol. 53, No. 6. Pergamon, 980-1023, (2008), doi: 10.1016/j.pmatsci.2008.05.001.

Persian Abstract

چکیده

در سی سال گذشته، فرآیند جوش‌کاری اصطکاکی اغتشاشی (FSW) به دلیل نتایج رضایت‌بخش حاصل از تغییر شکل شدید و گرمای کم ورودی در طول تولید اتصال جوش داده شده، اهمیت قابل توجهی یافته است. این عناصر برای پیاده سازی FSW در سیستم‌های جوش داده شده مختلف از جمله اتصالات آلومینیوم – فولاد در نظر گرفته شده‌اند. در این اتصالات غیرمشابه، هدف اصلی به دست آوردن یک اتصال جوش‌داده‌شده با رفتار مکانیکی قابل قبول بود. اخیراً برخی از مقاله‌ها بر درک فرآیند اتصالات غیرمشابه، عمدتا در جریان فلز و واکنش آن در برابر خوردگی متمرکز شده‌اند. با این حال، در اتصالات فولاد–آلومینیم، حضور ذرات فولاد در منطقه‌ی گرده‌ی جوش یک امر عادی است. این رفتار مکانیکی و شیمیایی اتصالات جوش داده شده را تغییر می‌دهد. بنابراین، هدف این پژوهش ارزیابی سازوکارهای تولید این ذرات و تعیین اثر انحراف ابزار در شکل گیری آنها و پیشنهاد رفتار جریان مواد با استفاده از ذرات جدا شده به عنوان ردیاب است. مشخص شد که انحراف ابزار سیالیت فلز را کنترل می کند، که به دلیل کاهش بی‌نظمی در رابط فولاد–آلومینیم امکان تجمع ذرات فولادی در سمت جلو و کاهش مقدار آن را فراهم می کند. به همین ترتیب، همان طور که در اتصالات آلومینیوم ذکر شده است، جریان فلز در سمت عقب مشاهده شد. در مقابل، در سمت جلو، یک حرکت برشی، فشار به سمت پایین و حرکت جانبی به سمت عقب، که توسط مقاومت نیروی آهنگری بالای فلز و محدودیت‌های اعمال شده توسط فولاد و ورق پشت‌بند حرکت می کند، وجود دارد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 04, (April 2021) 1035-1045

International Journal of Engineering

Journal Homepage: www.ije.ir

Experimental Study of the Effect of a Non-oxygenated Additive on Spark-ignition Engine Performance and Pollutant Emissions

S. Davari[a], F. Ommi*[a], Z. Saboohi[b], M. Safar[a]

[a] *Department of Mechanical Engineering, Tarbiat Modares University, Tehran, Iran*
[b] *Aerospace Research Institute (ARI), Ministry of Science, Research, and Technology, Tehran, Iran*

PAPER INFO

Paper history:
Received 23 November 2020
Received in revised form 19 January 2021
Accepted 02 February 2021

Keywords:
Fuel Additive
Pollutant Emissions
Engine Performance
Four-stroke Spark Ignition Engine
Spark Plugs

ABSTRACT

In this study, effects of adding butene, homopolymer to gasoline on the performance of a four-stroke spark ignition (SI) engine and pollutant emissions have been investigated. This additive increases the octane number of gasoline. In this research, the additive was combined with a non-leaded gasoline. Also, in addition to fuel changes and the use of additives, engine spark plugs were replaced and three types of spark plugs were used for this study. These include single electrode spark plug, dual electrode spark plug and Platinum+4 spark plug. The results of experimental tests showed that with the addition of additive to gasoline, the brake torque and braking power were increased with the use of each of the three spark plug type. The results revealed that by combining gasoline and additive, carbon dioxide and nitrogen oxides emissions were increased. On the other hand, unburned hydrocarbons and carbon monoxide emissions from the engine were reduced. In addition, it was concluded that changing the spark plugs had slight effect on engine performance and pollutant emissions, and that the results of experimental tests using all three types of spark plugs were almost identical.

***doi**: 10.5829/ije.2021.34.04a.30*

1. INTRODUCTION

Nowadays, internal combustion engines play an important role in human social life. In addition to the automotive industry, different types of internal combustion engines are widely used in other industries such as aviation, power generation, petroleum industries, etc. The limited availability of fossil fuels on the one hand and the pollutant emissions restriction laws on the other hand, have led researchers to move continuously towards optimization in the design of internal combustion engines as well as the development of new fuels with the highest efficiency and lowest levels of emissions [1].

The use of fuel additives or the improvement of combustion properties of fuel has always been one of the important issues in improving the performance of internal combustion engines. Tetraethyl lead (TEL), with the molecular formula $((CH_3CH_2)_4)Pb$, was blended with gasoline in the early 1920s as an octane number enhancer. The burning of TEL causes the emission of lead oxide, which is a very toxic substance [2]. Methyl tert-butyl ether (MTBE) has been used in the United States since late 1970s in a small amount as an octane number enhancer. Most of the refineries have used this substance as an additive because of the desired structural properties of MTBE and its cost-effectiveness. MTBE is a cancerous substance known to humans and animals, due to its solubility in groundwater [3]. Stratiev and Kirilov [4] investigated the effect of adding ferrocene ($Fe(C_5H_5)_2$) to gasoline as an octane enhancer and concluded that the motor octane number (MON) increased by 2.7 units, and the research octane number (RON) increased by 5.7 units. Ferrocene is known as an inexpensive additive, therefore the refineries replaced this substance with tetraethyl lead. One of the disadvantages of using ferrous as an additive is the formation of iron deposits. These deposits form on a

Corresponding Author Institutional Email: *fommi@modares.ac.ir* (F. Ommi)

Please cite this article as: S. Davari, F. Ommi, Z. Saboohi, M. Safar, Experimental Study of the Effect of a Non-oxygenated Additive on Spark-ignition Engine Performance and Pollutant Emissions, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1035-1045

spark plug and produce a conductive surface. Rashid et al. [5] experimentally studied the effect of three gasoline fuels (RON 95, 97 and 102) on engine performance and emission of hazardous gases. They stated that increasing the fuel octane number, gives rise to increase the amount of power and torque while it decreases the amount of NO_X emissions. Ebrahimi and Mercier [6] investigated the effect of simultaneous use of gasoline and natural gas on the performance of the SI engine. He concluded that the use of natural gas increased the specific fuel consumption and the thermal efficiency. On the other hand, brake torque, exhaust gas temperature and lubricante oil temperature were reduced compared to gasoline.

Oxygenated additives have been abundantly used in fuels. Pan et al. [7] investigated the effect of 2-phenylethanol on gasoline. 2-Phenylethanol is an additive with a high octane number (about 110) that is derived from biofuel. They observed that if this additive was added, the pressure of the cylinder and the released heat decrease. They also found that, adding this additive also reduces combustion time and ignition delay. According to their results, 2-phenylethanol has a very good resistance to engine knock, due to its high octane number and its chemical structure. They concluded that the addition of this additive to gasoline leads to a reduction in nitrogen oxides, while it gives rise to increase the amount of unburned hydrocarbons and carbon dioxide. Okoronkwo et al. [8] used petrol–ethanol–diethyl-ether blend with various percentage proportion in a spark ignition (SI) engine to reduce carbon monoxide emissions. Their results showed that by increasing the amount of oxygen in the fuel, the amount of carbon monoxide emissions reduces. Srinivasan et al. [9] investigated the effect of ethanol–gasoline blend with additives on a multi-cylinder spark ignition engine. The fuel additive used by them included a mixture of toluene, methanol, isopropyl alcohol, acetone, and xylene. They concluded that the brake thermal efficiency increases with the use of alternative fuels. Also, their results showed that the amount of CO, CO_2, HC, and NOx emissions appreciably decreased. Zaharin et al. [10] examined the effect of iso-butanol additives in ethanol-gasoline blend on fuel properties, performance and emission characteristics of a four-cylinder spark-ignition engine. They determined that the blended fuels displayed higher brake power than pure gasoline. Also, significant reduction in brake specific fuel consumption obtained compared to base gasoline fuel. In addition, they found that with these blended fuel samples, the exhaust gas temperature increased. According to their research, carbon monoxide and unburned hydrocarbons emissions were reduced. In contrast, carbon dioxide and nitrogen oxides emissions increased. Zamankhan et al. [11] and Valihesari et al. [12] used a new combination of fuel with oxygenate additives and metal nanoparticles in a spark ignition engine to study the effects of the novel blends on the engine parameters and also the amount of emissions. Based on their results, the performance parameters of the engine were improved and the amount of pollutants was reduced. Amirabedia et al. [13] investigated the effect of adding ethanol and Mn_2O_3 and Co_3O_4 nano additives to pure gasoline on the performance parameters and pollutant emissions of a spark ignition engine. They concluded that a combination of 10% ethanol and 20ppm Mn_2O_3 was the best fuel blend.

Another group of fuel additives used in gasoline is the additives containing aromatic hydrocarbons. Patil et al. [14] examined the effect of adding aromatic hydrocarbons to gasoline and concluded that adding aromatic hydrocarbons to gasoline increases RON and MON simultaneously. While the effect of Alkylbenzene on MON is more than RON. Also, with the addition of aromatic alcohols to gasoline, the octane number was increased by more than 10%. Demirbas et al. [3] used a catalyst in the internal combustion engine to reshape the structure of hydrocarbons and convert them into ring compounds. This catalyst converts alkanes with direct chain structures into cycloalkanes, and then converted them into aromatic compounds that have high octane numbers. They studied the effect of the naphtha catalyst on gasoline, in which 60% of saturated hydrocarbons with a low octane number turned into aromatic compounds with higher octane numbers.

As discussed, the use of innovative compounds in fuel additives is a broad field. Using fuel additives is one of the main ways to ameliorate engine performance, and decrease pollutant emissions. In this study, the effect of a non-oxygenated additive has been studied. In this way, the effect of fuel additive was tested using three different types of engine spark plugs. In this study, the new additive was examined with 15 percent of volumetric ratio. The aim of this study was to investigate changes in engine performance and pollutant emissions with varying fuels and spark plugs. In the following, the additive used, the laboratory equipment, as well as the experimental tests methodology are stated. Finally, the main results of this research are presented and analyzed.

2. MATERIALS AND METHODS

2. 1. Fuel Preparation The fuels used in this study include unleaded gasoline and gasoline with butene, homopolymer. Table 1 shows the values of octane number of fuels used in this study, which include one unleaded gasoline and another gasoline with the addition of additive.

2. 2. Engine and Experimental Setup Tests were performed in the engine and propulsion laboratory

TABLE 1. Values of octane number of fuels used in this study

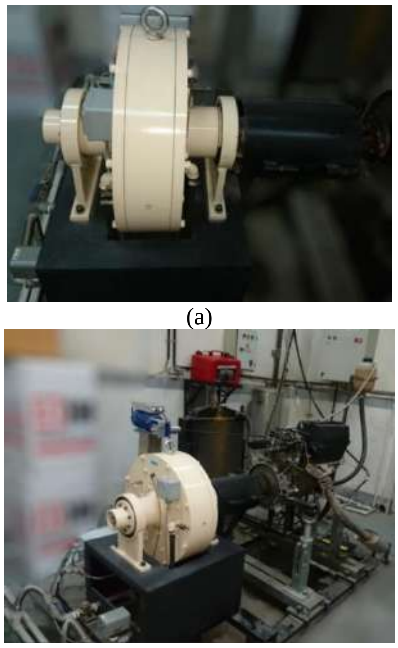
	Normal gasoline	Gasoline with additive
Motor Octane Number (MON)	83.6	92.1
Research Octane Number (RON)	87.3	`96.7
Antiknock Index (AKI)	85.45	94.4
Lower heating value (MJ/kg)	43.5	45.7
Latent Heat of Vaporization (kJ/kg)	350	331

at Tarbiat Modarres University employing a spark ignition engine model number XU7JP/L3. The specifications of the XU7JP/L3 engine are presented in Table 2. A 130kW eddy-current dynamometer (manufactured by MPA Company) was used to gauged power and torque of the engine. Figure 1 shows the rig used dynamometer and the engine test bed. The dynamometer was connected to the crankshaft through a rotating shaft, and by applying load to the engine, the engine power and torque were determined at each speed. The dynamometer was located inside a steel chamber to provide the necessary safety when its working. The load was applied to the engine using a computer dynamometer software. In this study, in addition to fuel changes and the use of additives, engine spark plugs were also replaced and three types of spark plugs were used for this research. These include single electrode spark plug, dual electrode spark plug and Platinum+4 spark plug. Figure 2 shows the three spark plugs used in the experiments. A cooling tower was used to prevent overheating of the dynamometer. In order to measure the pollutants, the CAP3200 analyzer, made by Capelec Company (France) was used. The pollutants measured by the analyzers include carbon dioxide, carbon monoxide, unburned hydrocarbons (UHCs) and nitrogen oxides. The analyzer is shown in Figure 3.

2. 3. Experimental Tests Methodology In this section, the steps of measuring the engine performance as well as the emission levels of the pollutants are discussed.

TABLE 2. The test engine characteristics

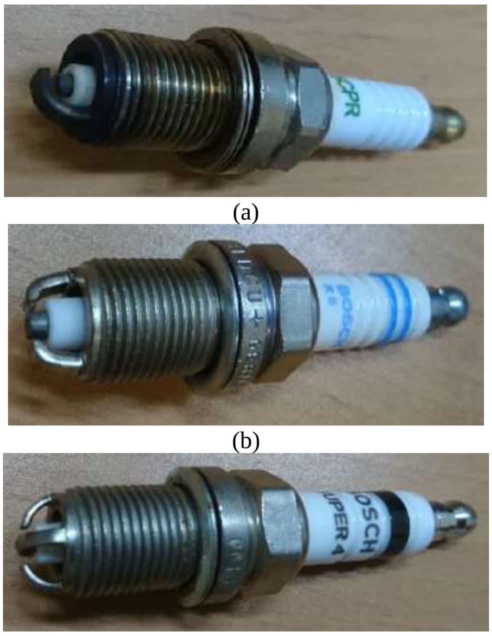
Company	Peugeot
Engine Model	XU7JP/L3
Number of cylinder	4
Capacity (cm^3)	1761
Compression ratio	10.2:1
Cylinder Bore (mm)	81.4
Stroke (mm)	83
Maximum Power (hp)	100 at 6000 rpm
Maximum Torque (Nm)	153 at 3000 rpm



(a)

(b)

Figure 1. View of (a) the dynamometer and (b) the engine test bed



(a)

(b)

(c)

Figure 2. Three types of used spark plug: (a) Single electrode spark plug, (b) Dual electrode spark plug, (c) Platinum+4 spark plug

In the laboratory test stand, the engine was initially turned on, and it operated with fixed rotational speed for10 minutes to reach steady state condition. The studied additive with a 15% volumetric ratio was combined with lead-free gasoline, and the resulting fuel and gasoline without additives were subjected to performance tests.

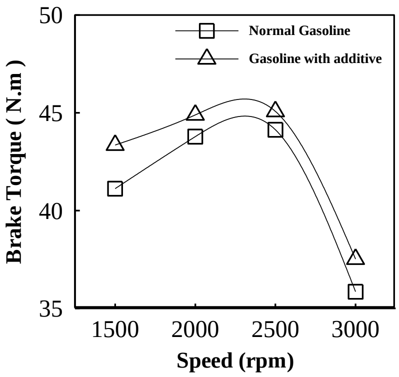
Figure 3. CAP3200 Analyzer

The tests were carried out in such a way that in order to investigate the impact of the additive in different engine conditions, the single electrode spark plug was first examined with both types of fuel. Then, in the second stage, the dual electrode spark plug was studied with both types of fuel, and finally, the platinum+4 spark plug was examined with both types of fuel. For each fuel mixture, the engine performance tests were performed for four different modes of the engine operational conditions at 1500, 2000, 2500, and 3000 (rpm). All tests were carried out with a fixed (50 percent Throttle) engine load, and the engine speed was automatically applied by the test stand software. During each run, the engine operated for 5 minutes to get a steady state operating condition. Power and torque of the engine were measured through the dynamometer and these data were automatically saved in a folder. The sensor of the analyzer was connected to the exhaust outlet and the amount of pollutants was obtained during the test and at different engine speed. At the end of each test, the sensor was cleaned to increase the accuracy of the measurement. Then, characteristics such as power, torque, fuel consumption, pollutant emissions, exhaust outlet temperature and pressure, oil temperature and pressure, etc., were compared for each fuel type and at different engine speed as functional parameters. Experimental tests were repeated three times to evaluate the repeatability of the test results, and the mean value was reported at the end. Also, in this research, at each stage of the test, the comparison of the results between ordinary gasoline and gasoline with the additive in a fixed engine speed was conducted. In order to validate the results, the experimental results were compared with those published by the engine manufacturer. The amount of throttle considered was equal to 50 percent in all graphs provided by the manufacturer. For this reason, in this experiment, the amount of throttle was considered constant (equal to 50 percent), and throttle is not a variable parameter. It should be noted that all experiments were repeated three times and the maximum error rate of the results was 0.5%. Also, all sources of experimental error were identified and eliminated as much as possible. These sources of error include emissions measuring errors, dynamometer errors, data averaging errors and etc. In this way, each of the experiments that were repeated three times were compared with each other, and then their error percentage was calculated. Another important issue was determining the volume fraction of the additive, which should be considered both in terms of improving engine performance and reducing pollutants, as well as in terms of cost. Therefore, this additive was studied and analyzed from all aspects of this study. The theoretical implications of improving engine performance as well as reducing emissions if using this additive, which has higher octane and LHV content than pure gasoline, were tested by experimental tests.

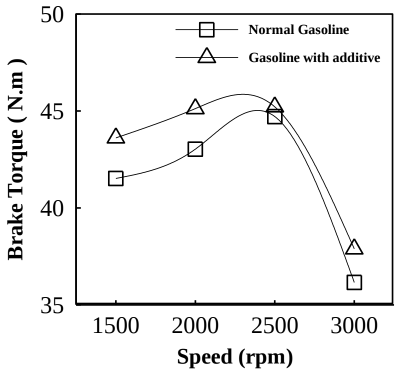
3. RESULTS AND DISCUSSION

3. 1. Engine Performance Parameters As shown in Figure 4, the addition of an additive to the fuel increases the brake torque, due to the process and the combustion cycle were completed, and the chemical energy of the fuel was more efficiently converted to mechanical work. It is also necessary to explain that with increasing engine speed to 2500 rpm, brake torque increases. This is due to the complete filling of the cylinder at engine breathing. However, at higher speeds due to less breathing time, the cylinder is not fully filled, and as a result the brake torque decreases.

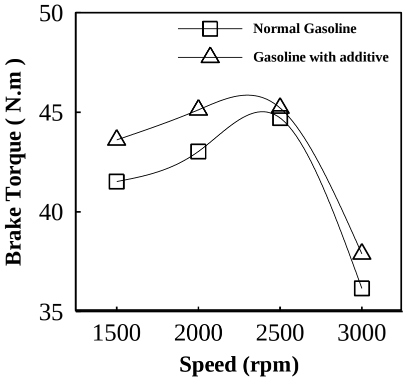
As shown in Figure 5, the brake power increases with the addition of the additive to the fuel, and this phenomenon is evident at all engine speeds. This can be attributed to the increase in octane number (As shown in Table 1) due to the mixing of fuel with additives, since the increase in octane number delayed the occurrence of combustion. Accordingly, the pressure in the piston compression stage decreases, and increases in its expansion stage. Consequently, the enclosed surface of the P-V curve increases and therefore the brake power increases. The brake power is directly related to the engine speed, hence as the engine speed increases the brake power increases as well. The heating value of the fuel increased by adding butene, homopolymer to gasoline and as result increases the engine power. As the engine speed increases, the effect of adding additive becomes more apparent, since fuel consumption increased and leads to enhanced engine power [15]. As presented in Table 1, the heating value of fuel with additive is more than that of normal gasoline. This means that the engine requires a higher amount of neat gasoline to generate the identical engine power as in fuel with additive. As a result, using fuel with butene, homopolymer increases power and torque compared to regular gasoline. Typically, engine power and torque are closely related to pressure inside cylinder. As shown in Figure 6, the fuel with additive displayed a higher pressure inside cylinder.



(a)



(b)

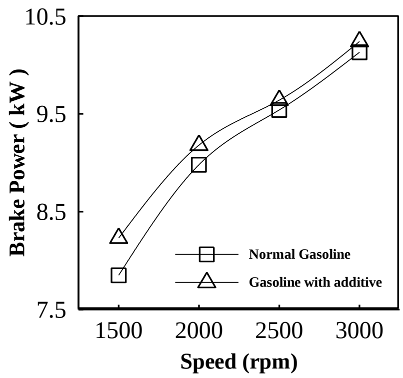


(c)

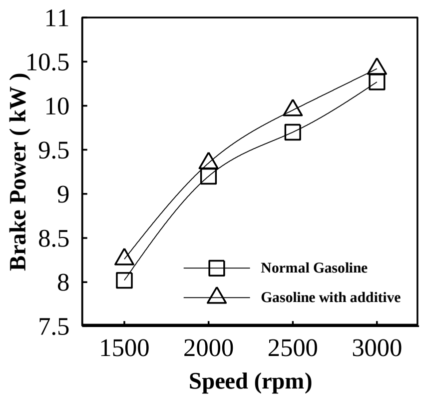
Figure 4. Experimental results of brake torque for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

Figure 6 shows the pressure inside cylinder changes against the crank angle for various fuel mixtures. Combustion chamber pressure is one of the most important parameters in determining the combustion process characteristics [16]. It can be seen that the pressure inside the chamber using gasoline with additive is slightly higher than pure gasoline. This is due to the lower heating value (LHV) of gasoline with butene, homopolymer is higher than conventional gasoline. Therefore, the pressure inside the cylinder is directly related to the engine power. Ozsezen and Canakci [17] derived that as the chamber pressure increases, the engine power increased.

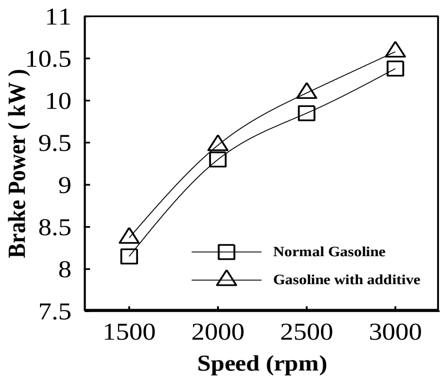
The percentage of changes in engine performance parameters for gasoline with the additive at different engine speeds is compared to conventional gasoline and data reported in Table 3.



(a)



(b)



(c)

Figure 5. Experimental results of brake power for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

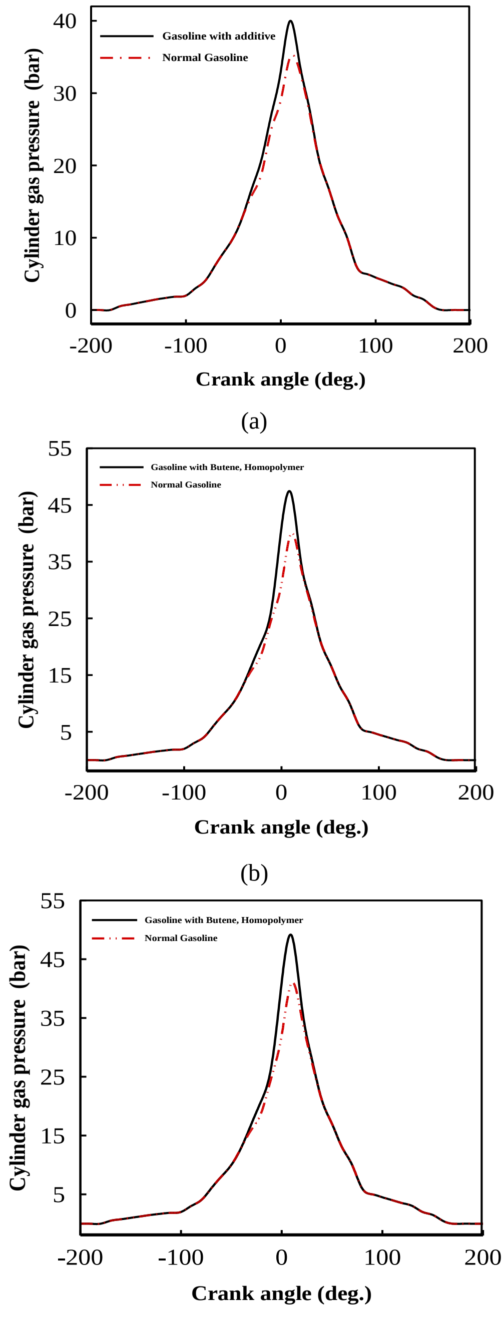


Figure 6. Effect of fuel blends on indicated pressure (IMEP) at engine speed 2500 rpm for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug

As can be seen, spark plug replacement has little effect on engine performance parameters and the performance of all three types is almost identical.

Figure 7 shows the ignition delay changes at different engine speeds for different types of spark plugs using pure gasoline. As shown in Figure 7, changing the spark plugs slightly reduces the ignition delay.

TABLE 3. Percentage increase of engine performance parameters when using gasoline with additive compared to conventional gasoline

Type of spark plug	Engine speed	Percentage increase of brake torque	Percentage increase of brake power
Single Electrode Spark Plug	1500	5.42	4.84
	2000	2.54	2.23
	2500	2.13	1.05
	3000	4.66	1.08
	Average Values	3.69	2.31
Dual Electrode Spark Plug	1500	4.94	2.99
	2000	4.77	1.63
	2500	0.94	2.58
	3000	5.11	1.46
	Average Values	3.94	2.16
Platinum+4 Spark Plug	1500	5.03	2.61
	2000	4.85	1.70
	2500	1.92	2.42
	3000	4.78	1.90
	Average Values	4.14	2.14

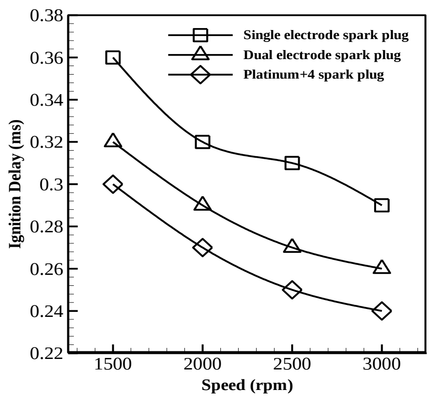
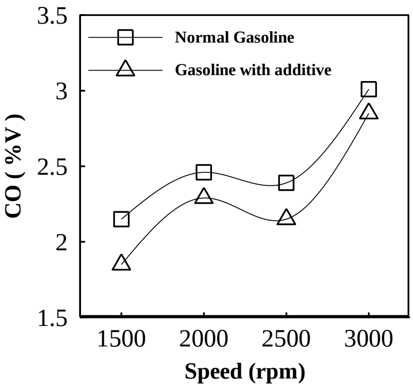


Figure 7. Variations of ignition delay according to different engine speeds for different types of spark plugs using pure gasoline

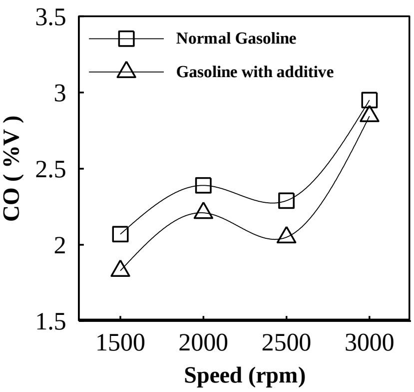
3. 2. Engine Exhaust Emissions

In this section, the effects of the addition of fuel additive on pollutant emissions were investigated. As shown in Figure 8, the amount of carbon monoxide released by the engine decreases by adding the additive to gasoline. The main contributor to controlling carbon monoxide is the excess air ratio, which was considered to be fixed in this study. As it was stated, the combustion process improves with the use of this additive. This means that the fuel burns more slowly and more time is required to complete the combustion reaction. Consequently, carbon monoxide emissions are reduced. Also, as the engine speed rises, the amount of carbon monoxide emissions increases. The reason is that, by increasing the engine speed, combustion reactions do not have enough time to form, and therefore incomplete combustion occurs. The main cause of carbon monoxide emissions is incomplete combustion [10]. In Figure 8, at 2000 rpm, the amount of carbon monoxide emissions first increased and then decreased at 2500 rpm (the opposite is shown in Figure 9 for carbon dioxide emissions) due to combustion fluctuations in the combustion chamber that dumps over time.

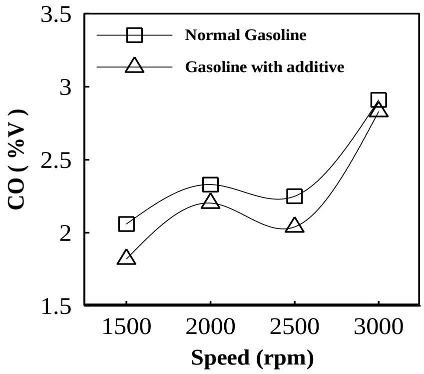
The effect of adding additive to gasoline on the amount of carbon dioxide emissions released from the engine is shown in Figure 9. The results showed that carbon dioxide emissions increased by adding additive to gasoline. As previously mentioned, due to the combustion improvement and the completion of the combustion process, the amount of carbon dioxide increases and the amount of carbon monoxide decreases.



(a)



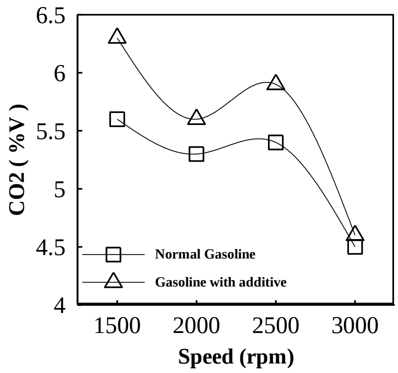
(b)



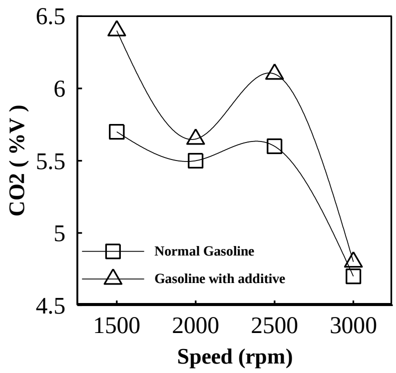
(C)

Figure 8. Experimental results of carbon monoxide for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

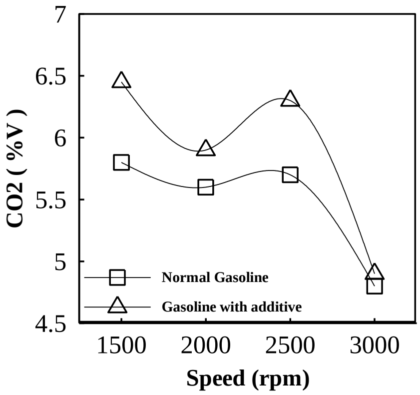
It is also apperceived that with increasing engine speeds, carbon dioxide emissions are reduced due to incomplete combustion. The reason that there is less time for combustion reactions in the engine. Carbon dioxide emissions display various conduct than carbon monoxide emissions [18].



(a)



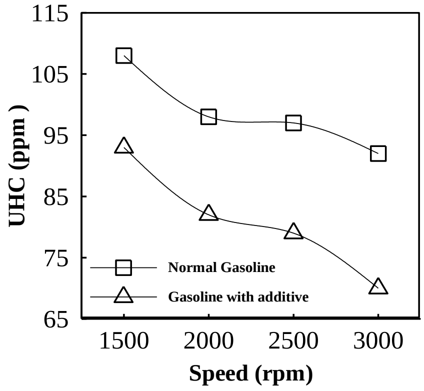
(b)



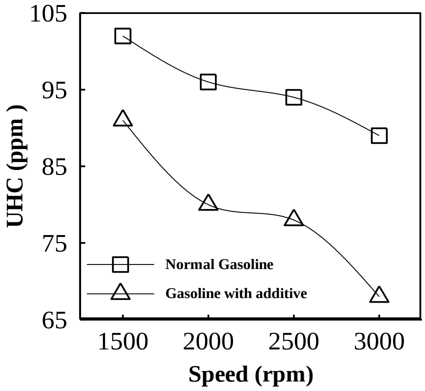
(c)

Figure 9. Experimental results of carbon dioxide for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

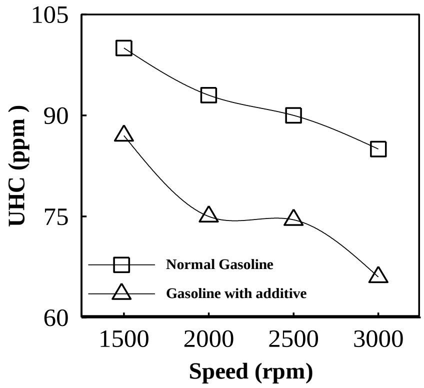
The effect of adding additive to gasoline on the amount of unburned hydrocarbons (UHCs) emissions produced by combustion is shown in Figure 10. The amount of UHC emissions increased if the fuel is not burned or incomplete combustion inside the chamber [10]. As can be seen from the figure, the increase in engine speed reduces the amount of unburned hydrocarbons emissions. As the engine speed rises, the opportunity for heat transfer from the engine is reduced, and thus the temperature of the combustion chamber wall increases, which consequently leads to decrease the thickness of the flame silencing layer near the wall. By adding the additive to the fuel, the amount of unburned hydrocarbons emissions decreases, The reasons for this decrease are: 1) Increasing the temperature of the combustion chamber, which causes the flame silencing to be delayed when it reaches the cylinder wall. This is because one of the main causes of the formation of unburned hydrocarbons emissions is the occurrence of flame silencing phenomenon in the adjacent of the cylinder wall. 2) Increasing the temperature of the combustion gases, which causes post-reactivity in exhaust gases and these hydrocarbons are oxidized at high temperatures. The fuel containing the additive has a higher octane number than conventional gasoline. The high octane number is one of the factors that increases the velocity of the flame. As a result, the flame does not turn off when it reaches the grooves inside the cylinder and unburned hydrocarbons emissions are reduced. UHC emissions were more reduced than other emissions. The UHC emissions decreased with the addition of the butene, homopolymer to the fuel. Increasing the temperature of the combustion gases caused post-reactivity in the exhaust gases. UHC emissions are more sensitive to temperature than other emissions. As a result, UHC emissions were decreased with increasing temperature.



(a)



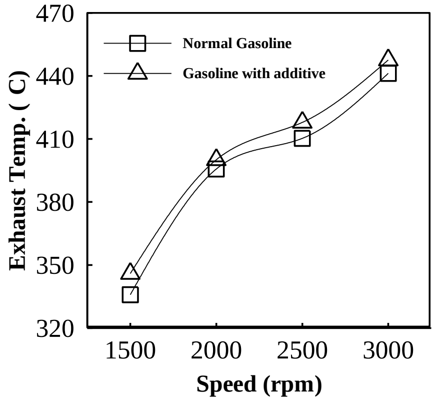
(b)



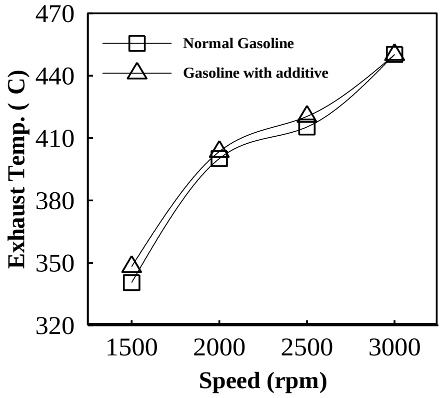
(c)

Figure 10. Experimental results of unburned hydrocarbons for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

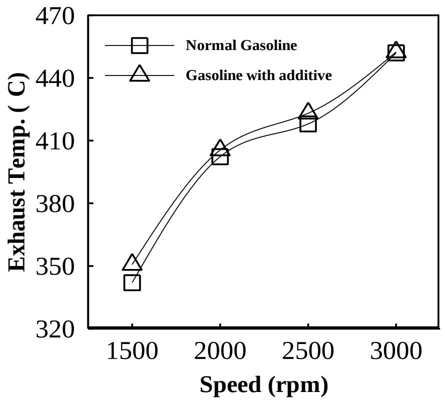
Figure 11 shows the exhaust gas temperature (EGT) trend for different fuels. EGT changes proportionally with the maximum cylinder temperature [19]. Exhaust gas temperature is higher when using gasoline with butene, homopolymer. Pure gasoline has greater latent heat than fuel with additive, as stated formerly. As a result, vaporization of conventional gasoline reasons a higher temperature drop in combustion chamber.



(a)



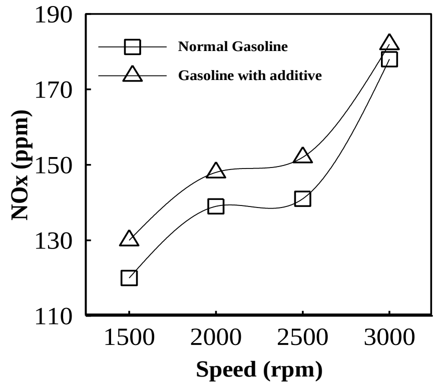
(b)



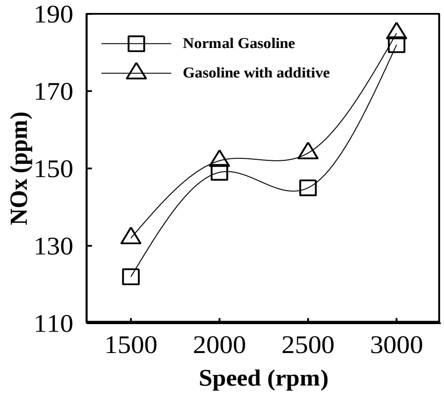
(c)

Figure 11. Experimental results of temperature of the exhaust gases for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds

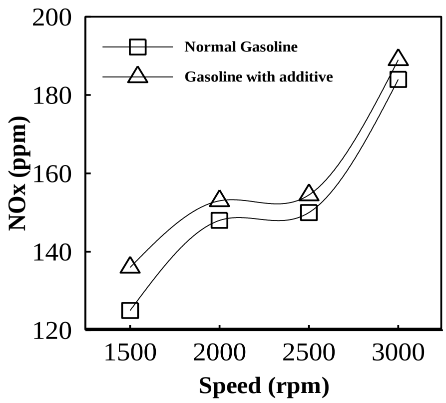
The effect of adding additive to gasoline on the amount of nitrogen oxides emissions produced by combustion is illustrated in Figure 12. The formation of NOx emissions is influenced by the combustion temperature [20]. As the engine speed increases and the additive is used, the amount of nitrogen oxides emissions increases, because at high engine speeds the number of work cycles increases over a given time, and as a result the heat transfer decreases and the maximum combustion temperature in the cylinder increases. It should be noted that the highest amount of nitrogen oxides emissions produced in the internal combustion engines is from the thermal path (Zeldovich). The Zeldovich mechanism



(a)



(b)



(c)

Figue 12. Experimental results of nitrogen oxides for (a) Engine with single electrode spark plug, (b) Engine with dual electrode spark plug, (c) Engine with platinum+4 spark plug at different fuel blends and engine speeds explicates the formation of thermal NOx [21]. Therefore, the production rate of this pollutant has a direct relation with the highest combustion temperature. The thermal path of nitrogen oxide production increases exponentially with increasing temperature, and is very sensitive to temperature. In other words, nitrogen and oxygen react at high temperatures, so the high temperature is the significant factor for the formation of nitrogen oxides emissions [22]. Using fuel with the additive, more complete combustion is formed inside the cylinder, which increases the temperature of the cylinder by complete combustion, and therefore the amount of nitrogen oxides emissions increases as temperature rises. NOx emissions increased with increasing engine speed due to increased temperature and pressure inside the combustion chamber. Therefore, NOx emissions are directly related to the temperature of the combustion chamber [16, 23, 24].

Percentage increase or decrease of pollutants in the case of using gasoline with the additive compared to the use of ordinary gasoline in different engine speed is reported in Table 4.

TABLE 4. Percentage changes in engine exhaust emissions when using gasoline with additive compared to conventional gasoline

Type of spark plug	Engine speed	Percentage reduction of CO	Percentage increase of CO_2	Percentage reduction of UHCs	Percentage increase of NO_X
Single Electrode Spark Plug	1500	13.95	12.5	13.88	8.33
	2000	6.91	5.66	16.33	6.47
	2500	10.04	9.26	18.56	7.80
	3000	5.31	2.22	23.91	2.25
	Average Values	9.05	7.41	18.17	6.21
Dual Electrode Spark Plug	1500	11.59	12.28	10.78	8.19
	2000	7.53	2.72	16.66	2.01
	2500	10.48	8.93	17.02	6.21
	3000	3.56	2.13	23.59	1.65
	Average Values	8.29	6.52	17.02	4.52
Platinum +4 Spark Plug	1500	11.65	11.21	13.01	8.80
	2000	5.45	5.36	19.35	3.38
	2500	9.33	10.53	17.22	3.00
	3000	2.75	2.08	22.35	2.71
	Average Values	7.29	7.29	17.98	4.47

4. CONCLUSIONS

The use of fuel additives is one of the main strategies for improving engine performance and reducing the amount of pollutants. In this research, the effects of adding a non-oxygenated additive to gasoline on engine performance and pollutant emissions was investigated. Also, engine spark plugs were replaced and tests were repeated using each of the spark plugs. The results of experimental tests showed that by adding the butene, homopolymer to gasoline, brake torque, and brake power were increased by an average of 4% and 2%, respectively. Also, with by increasing the engine speed up to 2500 rpm, the brake torque was increased, while its amount was decreased at higher speeds. On the other hand, brake power was increased with increasing engine speed.

By using the additive, the amount of carbon monoxide (CO) and unburned hydrocarbons (UHCs) emissions were decreased by an average of 8% and 18%, respectively. Also, with increasing engine speed, the carbon monoxide emission was increased, but on the other hand, the amount of unburned hydrocarbons emission was decreased.

By adding the additive to fuel, carbon dioxide (CO_2) and nitrogen oxides (NOx) emissions were increased by an average of 7% and 6%, respectively. Also, the results revealed that with increasing engine speed, carbon dioxide emission was decreased, but on the contrary, nitrogen oxides emission was increased.

As previously mentioned, in this study, three types of spark plugs were used, which include single electrode, dual electrode and Platinum+4. As a result, according to experimental tests, spark plug replacement had little effect on the values of performance parameters and pollutant emissions of engine, and their values remained almost constant. The effect of multi-electrode spark is long-term and the short-term effect is not significant. Because erosion occurs over time and ignition by the spark plug is not done well, on the other hand, multi-electrode spark plugs are more resistant to this phenomenon and therefore have better performance in a long time Operation.

5. REFERENCES

1. Huang, Y.-H. and Wu, J.-H., "Analysis of biodiesel promotion in taiwan", ***Renewable and Sustainable Energy Reviews***, Vol. 12, No. 4, (2008), 1176-1186. doi: 10.1016/j.rser.2007.01.009
2. Wiesenthal, T., Leduc, G., Christidis, P., Schade, B., Pelkmans, L., Govaerts, L. and Georgopoulos, P., "Biofuel support policies in europe: Lessons learnt for the long way ahead", ***Renewable and Sustainable Energy Reviews***, Vol. 13, No. 4, (2009), 789-800. doi: 10.1016/j.rser.2008.01.011
3. Demirbas, A., Balubaid, M., Basahel, A., Ahmad, W. and Sheikh, M., "Octane rating of gasoline and octane booster additives", ***Petroleum Science and Technology***, Vol. 33, No. 11, (2015), 1190-1197. doi: 10.1080/10916466.2015.1050506

4. Stratiev, D. and Kirilov, K., "Opportunities for gasoline octane increase by use of iron containing octane booster", ***Petroleum & Coal***, Vol. 51, No. 4, (2009), 244-248.

5. Rashid, A.K., Abu Mansor, M.R., Ghopa, W.A.W., Harun, Z. and Mahmood, W.M.F.W., "An experimental study of the performance and emissions of spark ignition gasoline engine", ***International Journal of Automotive & Mechanical Engineering***, Vol. 13, No. 3, (2016). doi: 10.15282/ijame.13.3.2016.1.0291

6. Ebrahimi, R. and Mercier, M., "Experimental study of performance of spark ignition engine with gasoline and natural gas", Vol. 24, No. 1, (2011), 65-74.

7. Pan, M., Wei, H., Feng, D., Pan, J., Huang, R. and Liao, J., "Experimental study on combustion characteristics and emission performance of 2-phenylethanol addition in a downsized gasoline engine", ***Energy***, Vol. 163, (2018), 894-904. doi: 10.1016/j.energy.2018.08.130

8. Okoronkwo, A., Igbokwe, J., Ezurike, B. and Oguoma, O., "The emission characteristics of a petrol–ethanol–diethyl-ether blend as a carbon monoxide reduction additive in spark ignition engine", ***International Journal of Ambient Energy***, Vol. 39, No. 4, (2018), 360-364. doi: 10.1080/01430750.2017.1303627

9. Srinivasan, C.A. and Saravanan, C., "Emission reduction on ethanol–gasoline blend using fuel additives for an si engine", ***Energy Sources, Part A: Recovery, Utilization, and Environmental Effects***, Vol. 35, No. 12, (2013), 1093-1101. doi:10.1080/15567036.2011.584114

10. Zaharin, M.S.M., Abdullah, N.R., Masjuki, H.H., Ali, O.M., Najafi, G. and Yusaf, T., "Evaluation on physicochemical properties of iso-butanol additives in ethanol-gasoline blend on performance and emission characteristics of a spark-ignition engine", ***Applied Thermal Engineering***, Vol. 144, (2018), 960-971. doi: 10.1016/j.applthermaleng.2018.08.057

11. Zamankhan, F., Pirouzfar, V., Ommi, F. and Valihesari, M., "Investigating the effect of mgo and ceo 2 metal nanoparticle on the gasoline fuel properties: Empirical modeling and process optimization by surface methodology", ***Environmental Science and Pollution Research***, Vol. 25, No. 23, (2018), 22889-22902.

12. Valihesari, M., Pirouzfar, V., Ommi, F. and Zamankhan, F., "Investigating the effect of fe2o3 and tio2 nanoparticle and engine variables on the gasoline engine performance through statistical analysis", ***Fuel***, Vol. 254, (2019), 115618. doi: 10.1016/j.applthermaleng.2018.08.057

13. Amirabedia, M., Jafarmadar, S., Khalilarya, S. and Kheyrollahi, J., "Experimental comparison the effect of mn2o3 and co3o4 nano additives on the performance and emission of si gasoline fueled with mixture of ethanol and gasoline", ***International Journal of Engineering***, Vol. 32, No. 5, (2019), 769-776. doi: 10.5829/ije.2019.32.05b.19

14. Patil, A.R., Yerrawar, R., Nigade, S., Chavan, O., Rathod, H. and Hiran, B., "Literature review on need of composite additives for si engine", ***International Journal for Research and Development in Technology***, Vol. 2, (2014), 8-12.

15. He, B.-Q., Liu, M.-B. and Zhao, H., "Comparison of combustion characteristics of n-butanol/ethanol–gasoline blends in a hcci engine", ***Energy Conversion and Management***, Vol. 95, (2015), 101-109. doi: 10.1016/j.enconman.2015.02.019

16. Thangavel, V., Momula, S.Y., Gosala, D.B. and Asvathanarayanan, R., "Experimental studies on simultaneous injection of ethanol–gasoline and n-butanol–gasoline in the intake port of a four stroke si engine", ***Renewable Energy***, Vol. 91, (2016), 347-360. doi: 10.1016/j.renene.2016.01.074

17. Ozsezen, A.N. and Canakci, M., "Performance and combustion characteristics of alcohol–gasoline blends at wide-open throttle", ***Energy***, Vol. 36, No. 5, (2011), 2747-2752. doi: 10.1016/j.energy.2011.02.014

18. Li, Y., Ning, Z., Chia-fon, F.L., Yan, J. and Lee, T.H., "Effect of acetone-butanol-ethanol (abe)–gasoline blends on regulated and unregulated emissions in spark-ignition engine", ***Energy***, Vol. 168, (2019), 1157-1167. doi: 10.1016/j.energy.2018.12.022

19. Eyidogan, M., Ozsezen, A.N., Canakci, M. and Turkcan, A., "Impact of alcohol–gasoline fuel blends on the performance and combustion characteristics of an si engine", ***Fuel***, Vol. 89, No. 10, (2010), 2713-2720. doi: 10.1016/j.fuel.2010.01.032

20. Deng, B., Fu, J., Zhang, D., Yang, J., Feng, R., Liu, J., Li, K. and Liu, X., "The heat release analysis of bio-butanol/gasoline blends on a high speed si (spark ignition) engine", ***Energy***, Vol. 60, (2013), 230-241. doi: 10.1016/j.energy.2013.07.055

21. Miller, J.A. and Bowman, C.T., "Mechanism and modeling of nitrogen chemistry in combustion", ***Progress in Energy and Combustion Science***, Vol. 15, No. 4, (1989), 287-338. doi: 10.1016/0360-1285(89)90017-8

22. Varatharajan, K. and Cheralathan, M., "Influence of fuel properties and composition on nox emissions from biodiesel powered diesel engines: A review", ***Renewable and Sustainable Energy Reviews***, Vol. 16, No. 6, (2012), 3702-3710. doi: 10.1016/j.rser.2012.03.056

23. Gonca, G., "Influences of different fuel kinds and engine design parameters on the performance characteristics and no formation of a spark ignition (si) engine", ***Applied Thermal Engineering***, Vol. 127, (2017), 194-202. doi: 10.1016/j.applthermaleng.2017.08.002

24. Elfasakhany, A., "Experimental investigation on si engine using gasoline and a hybrid iso-butanol/gasoline fuel", ***Energy Conversion and Management***, Vol. 95, (2015), 398-405. doi: 10.1016/j.enconman.2015.02.022

Persian Abstract

چکیده

در این پژوهش، تاثیر افزودن بوتن، هموپلیمر به بنزین بر پارامترهای عملکردی و انتشار آلاینده‌ها از یک موتور اشتعال جرقه‌ای بررسی شده است. این افزودنی موجب افزایش عدد اکتان سوخت شد. همچنین در این پژوهش علاوه بر سوخت، شمع‌های موتور نیز تعویض گردید. شمع‌های مصرفی شامل شمع تک، دو و چهار الکترودی بودند. نتایج آزمون‌های تجربی نشان داد که با افزودن بوتن، هموپلیمر به سوخت، مقادیر گشتاور ترمزی، توان ترمزی، آلاینده دی اکسید کربن و آلاینده اکسیدهای نیتروژن با استفاده از تمامی شمع‌ها افزایش می‌یابند. همچنین در نقطه مقابل انتشار آلاینده‌های هیدروکربن‌های نسوخته و مونواکسید کربن کاهش یافت. علاوه بر این، این نتیجه حاصل شد که تغییر شمع‌ها تأثیر اندکی بر عملکرد موتور و انتشار آلاینده‌ها دارد و نتایج آزمون‌های تجربی آ با استفاده از هر سه نوع شمع تقریباً یکسان است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Simple Slope Stabilization on Quartz Sandstone using Horizontal Drain

S. Supandi*

Department of Mining Engineering, Institut Teknologi Nasional Yogyakarta, Indonesia

PAPER INFO

Paper history:
Received 15 November 2020
Received in revised form 31 January 2020
Accepted 08 February 2021

Keywords:
Depressurization
Dewatering
Hydrogeology
Slope
Stabilization
Horizontal Drain

ABSTRACT

Maintaining stability of public facilities, especially public roads, is very necessary, so the stabilization efforts must be carried out. The water from seepage that comes out of sandstone layer carries material with fine particles and this causes the rock cohesion to decrease and thus leading to scouring. As a result of the scouring, the layer becomes overstep and can disturb road stability. This study was conducted to provide simple but measurable recommendation for maintaining road stability after the previous stabilization effort failed. The method used a fluid mechanics approach in which water from the formation was given space to come out of the formation without carrying fine particles (cement). The analysis was carried out using the finite element method by installing horizontal drain pipe. The result of analysis shows that the horizontal drain installation helps water to come out without creating an overstep layer. The recommended horizontal drain is 8 m long with a slope of at least 3% and a length of 1 m that must enter the sandstone formation as a water source. With this method, the road stability can be maintained and the stabilization can be carried out in an easy, inexpensive, and applicable way.

doi*: 10.5829/ije.2021.34.04a.31*

NOMENCLATURE

σ'	Effective stress	Re	Reynolds number
σ_n	Normal stress	ρ	Water density
u	Pore pressure	v	Velocity of water flow
τ	Shear strength	D	Inner diameter of pipe
φ	Friction angle	μ	Viscosity of water
c	Cohesion	K	Hydraulic conductivity

1. INTRODUCTION

Slope stabilization is mandatory to maintain existing infrastructure around the location. The stabilization must be carried out using a way that is able to function optimally, which means that technically it is capable of maintaining stability but economically the cost is minimum. In a failure of stabilization effort on a slope adjacent to the public road, the scouring occurred on the slope which was led to overstep and disturbed the stability of the public road.

In Figure 1, the reddish brown and the dark gray materials are fill materials that are positioned above the in-situ materials which are mudstone and quartz sandstone units under the soil. On the channel wall, there is a layer of sand that continuously releases seepage throughout the year and leaves scouring in several places. The impact of scouring is the occurrence of overstep and from time to time, the scouring is getting bigger, that it will disturb the stability of public roads if it continues.

The previous research about the flow behavior has been conducted; it explained the vertical and horizontal seepages and their relationship to channel presence [1]. The groundwater will flow most rapidly through material with the greatest hydraulic conductivity to a place with the lower flow potential. In this case, the groundwater flow originating from the in-situ material will go to the exposed part through the fill material until it finally emerges as a spring. Based on the theory of hydromechanical coupling in soil and rock, the presence of groundwater has a direct impact on the material by

*Corresponding Author Email: supandi@sttnas.ac.id (S. Supandi)

Please cite this article as: S. Supandi, Simple Slope Stabilization on Quartz Sandstone using Horizontal Drain, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1046-1051

Figure 1. Diversion channel near access road in low water flow condition

increasing the volume of rock mass, especially in soil or pile material [2]. Indirectly, the changes in pore pressure will have an impact on changes in the mechanical properties of the material.

To overcome the existing hydrogeological condition, depressurization is required to increase the stability of material slope [3]. The effectiveness of horizontal drain installation is indicated by the increasing factor of safety (FoS) after the horizontal drain installation compared to the pre-installation condition. The horizontal drain installation should have been positioned as low as the location that is likely to be a slip surface [4]. In this study, the installed horizontal drain on fill material was not carried out by drilling, but by excavating and then backfilling after the casing had been installed.

The Installation of horizontal drain in slope stability aims to control and remove water from the material (desaturation) and depressurization [5]. When the groundwater is removed from the material, the ground water level will decrease; thus, resulting in a decrease in pore pressure in the material. The construction of vertical drain using column gravel will be reducing potential liquefaction or increasing pore pressure [6]. The design and success rate of this effort depends on local condition. The effectiveness of horizontal drain configuration includes the location, length, and spacing; slope geometry; as well as the characteristics of slope-forming material. The effectiveness of horizontal drain installation is indicated by the increasing factor of safety (FoS) after the horizontal drain installated [7].

The integration of geotechnical and hydrogeological aspects in slope stability analysis is very interesting but the implementation is still limited [8]. Water content has relationship with mechanical properties of rock and soil, where increasing water content will decrease mechanical properties of rock/soil [9-10]. This study aims to evaluate the stabilization effort using horizontal drain that was carried out in the study area to increase the effectiveness of the installed horizontal drain. This study will discuss the conditions before and after the installation of horizontal drain, and then compare their effects on slope stability.

The effective stress theory explains the behavior of soil and rocks in relation to the presence of groundwater (shown in Equation (1)) [11]. The effective stress (σ') is expressed as normal stress (σ_n) minus pore pressure (u) at a point. For geotechnical material, the strength of material is described by the Mohr-Coulomb law (shown in Equation (2)) as effective stress for a point along slip surface; where τ is shear strength at a point of interest, φ is friction angle, and c is cohesion.

$$\sigma' = \sigma_n - u \tag{1}$$

$$\tau = \mathrm{c} + (\sigma_n - u)\tan\varphi \tag{2}$$

Based on these two equations, it can be concluded that the greater the pore pressure in a material, the lower the strength of the material, and vice versa. Ideally, horizontal drain installation is positioned as low as possible at a location that is likely to be a slip surface. Given that hydraulic conductivity of filling material is much less than the conductivity / velocity of water flowing through a 3″ pipe in the same cross-sectional area, the flow in the pipe occurs freely without any over pressure.

The maximum velocity of flow in the pipe with free condition is calculated by the Reynolds number (Equation (3)) [12]. Basically, the seepage potential of in-situ material can be overcome by the layer of fill material. The stabilization effort using galam wood piles was carried out, but they could not cope with the movement of material.

$$Re = \frac{\rho \times v \times D}{\mu} \tag{3}$$

Diversion channel is an infrastructure designed to accommodate and drain water. Various studies on diversion channel have been conducted including physical dynamics of river and diversion channel due to water flow [13], implication of channel diversion [14], problem-solving method related to river diversion [15], design failure remediation [16], channel restoration, geometric design of channel [17], as well as case studies and reviews of river diversion in various locations [18].

The preliminary result analysis showed that wooden-pile installation can withstand material movement. In practice, creep of fill material occurred over time, causing design failure. Based on field investigation, the creep was identified and strongly believed, it is caused by groundwater flow (hydrogeology) factor in the fill material. This parameter has not been accommodated properly in preliminary analysis. For this reason, modeling and management of hydrogeological condition are required in the context of material stabilization.

2. MATERIALS AND METHODS

The research was conducted at one of the creek diversions in South Kalimantan, Indonesia, which is located right beside a public road. In general, the materials at research location are in the Warukin formation which is composed of sandstones, claystones, and coal with low hardness [19-22]. The sandstones are composed of clay-to-sand-sized quartz sandstones with brownish gray color [23]. The groundwater level is relatively high, about 10 m from the surface. The type of aquifer found is confined aquifer where the sandstone layer is bordered by claystone at the top and bottom. The seepage is found on the surface of sandstone slopes and water discharge is stable in either rainy or dry season. The water from the seepage that comes out all year round, than its mixed with very fine material.

The stabilization effort has been carried out by covering the sand layer with laterite material consisting of sand, clay, and gravel, so it covers the entire surface of the sand, and installing sheetpiles using galam wood. Both of these methods were unsuccessful because the laterite material deteriorated due to addition of water from seepage and the galam wood as a barrier was unable to hold it.

Based on the above conditions, the stabilization effort was carried out by keeping the seepage out of the formation without carrying rock cement so that soil cohesion can be maintained. A channel with perforated pipe was made for the water from the seepage to come out without carrying cement. The perforated pipe was made using PVC pipe in general and given a geotextile coating as a filter. In detail, the pipe is Schedule 120 PVC pipe which is 6″ in length with outer diameter of 6.625″ (141.3 mm) and inner diameter of 5.434″ (121.1 mm) and has collapse pressure of 370 psi (2551 kPa) (Figure 2).

The perforated pipe was only in the part that was being buried; however, at the outlet, it was solid. It was installed 1 m at the seepage, passing through the embankment zone before exiting the channel. The horizontal drain inclination is 5°, with the target of quartz sandstone unit contact. The total length of horizontal drain that has been installed, it is approximately 6 m, it is slightly entering the in-situ quartz sandstone material.

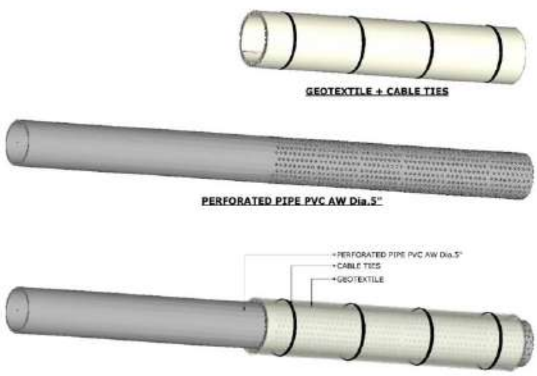


Figure 2. Design of perforated pipe for horizontal drain

The analysis was performed using combination limit equilbirum and finite element method with stratigraphy based on drilling data around the location. The geotechnical parameters were obtained based on laboratory analysis using triaxial and uniaxial methods with testing standard of ASTM. The determination of material properties on bedding contact referred in previous research in this area with value material properties is Cohesion 0 kPa, density 13 kg/cm^3 and friction angle 13 degree [24]. The failure was identified on the backfilling material only and determination bedding ratio was not applied [24].

3. RESULTS AND DISCUSSION

Based on field identification result, creeping only occurs in the fill material, while the landslides do not occur in the in-situ material. Therefore, the slope stability analysis was focused on the fill material. On surface of the embankment under fill material, there is sedimentation which is a mixture of slurry material and fill material on the channel bed that is also considered in the analysis. Probability analysis was performed based on the properties of fill material and sedimentation material (Table 1).

The values for hydraulic properties of in-situ material and embankment identified in the field are shown in Table 2. The in-situ material has a bedding that causes the vertical hydraulic conductivity to be smaller than the horizontal, which is normal to the bedding plane; for this reason, it is determined to have ratio of 1:10. Meanwhile, the soil material and embankment have a greater vertical conductivity ratio than in-situ material due to their loose condition, which is determined to be 1:1.

The hydrogeological analysis was carried out transiently. The stage of time was used 5 stages: 15 days

TABLE 1. Material statistics applied in probability analysis

Material name	Properties	Dist.	Mean	Min.	Max.
Fill material	Cohesion (kPa)	Normal	13.75	11.75	15.75
Fill material	Phi (°)	Normal	7	6	8
Fill material	Unit weight (kN/m^2)	Normal	17	16	18
Sediment	Cohesion (kPa)	Normal	1	0	2
Sediment	Phi (°)	Normal	7	6	8
Sediment	Unit weight (kN/m^2)	Normal	17	16	18

TABLE 2. Hydraulic conductivity of material

Material unit	Hydraulic conductivity	Kv/Kh	a (bedding angle)
Sandstone	8.25×10^{-05}	0.1	3.125
Mudstone	1.00×10^{-07}	0.1	3.125
Coal	3.65×10^{-06}	0.1	3.125
Fill material	4.00×10^{-04}	1	-
Soil	4.00×10^{-05}	1	-
Sedimentation	1.95×10^{-07}	1	-

(over 2 weeks), 30 days (1 month), 45 days (1.5 months), 60 days (2 months), and 90 days (3 months). The determination of stage is intended to know the short and long term variation of the effectiveness of installed horizontal drain.

3. 1. Horizontal Drain In the analysis, the length of horizontal drain was simulated with a length of 4 m, 8 m, and 12 m to the quartz sandstone with the same output position (stick out). The water flow in horizontal drain is considered to be free flow (laminar/non-turbulent) with Re ≤ 2000. By applying Equation (3), if the viscosity of water (μ) = 10^{-3} kg/(m.s), water density (ρ) = 1000 kg/m^3, and inner diameter of the Pipe 6" (D)= 0.1211 m, then the hydraulic conductivity is:

$$v=\frac{Re\times\mu}{\rho\times D}=\frac{2000\times10^{-3}\,\mathrm{kg/(m.s)}}{1000\,\mathrm{kg/m^3}\times0.1211\,\mathrm{m}}=0.0165\,\mathrm{m/s}$$

$$K\approx v=0.0165\,\mathrm{m/s}=1.65\times10^{-2}\mathrm{m/s}$$

Based on the above calculation, the hydraulic conductivity of horizontal drain is 1.65×10^{-2}m/s. This value is an approximation of the maximum value of the horizontal drain hydraulic conductivity used in the hydrogeological analysis.

3. 2. Innitial Condition Before Horizontal Drain Instalation The back analysis was carried out in initial condition before the horizontal drain installation. The analysis was carried out as an initial basis and validation of analysis before continuing the next analysis. The landslides occurred in the initial condition, but it did not occur immediately. This indicates that the value of stability factor was in a critical condition (FoS approaches 1.0), with a relatively high probability factor.

The back analysis was carried out in the initial condition at the stage of 15-90 days with stability factor value of 1.013-1006 and failure probability of 39.6-44.73%. Over time, there was saturation in the material which was caused by the decrease in the slope stability, that was accompanied by an increase in the probability of failure. This condition is considered to reflect the condition of the embankment, where the landslides did not occur at once but slowly creeping. Figure 3 shows the result of back analysis on stage of 90 days. The surface of fill material slope is a groundwater accumulation zone, where the flow in that part is the greatest. This is in accordance with the result of field identification that has been carried out that the groundwater condition greatly affects the stability of fill material.

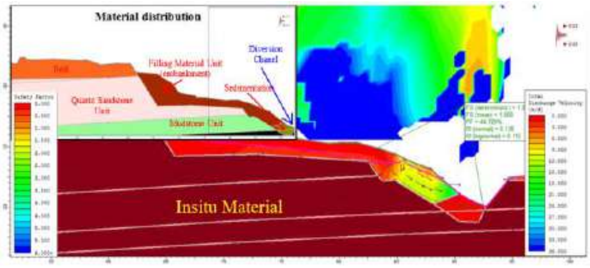


Figure 3. Back analysis on stage of 90 days showing total discharge velocity and accumulation on the slope surface of the fill material embankment. Insert shows material boundaries distribution

The series of analysis was carried out by the means of limit equilibrium and the failure probability analyses integrated with transient groundwater finite element analysis on the criteria that has been mentioned. Based on the results of the analysis, it is known that the most significant change in groundwater level in the study area occurred in the first 30 days. After 30 days, the changes in groundwater condition still occurred limitedly, but the changes were not so significant.

The correlation between pore pressure and time stage at query points located in the in-situ material contact with the fill material shows the most significant change in pore pressure from 15 days stage to 30 days stage. It was the case for all horizontal drain lengths that were analyzed. For the next stage, a decline in pore pressure value was identified but very small. Thus, the most striking changes in stability from the value of either stability factor or failure probability can be only observed up to the 30 days stage (Tables 3 and 4). It shows that the optimum effectiveness of horizontal drain installation in the study area was achieved after the 30 days stage.

Figure 4 shows the correlation between pore pressure and query point distance along a query line for

TABLE 3. Summary of limit equilibrium analysis

Horizontal drain length	FoS in stage of				
	15 days	30 days	45 days	60 days	90 days
Initial (none)	1.013	1.007	1.007	1.007	1.006
4 m	1.037	1.037	1.037	1.037	1.037
8 m	1.093	1.099	1.099	1.099	1.099
12 m	1.100	1.100	1.100	1.100	1.100

TABLE 4. Summary of failure probability analysis

Horizontal drain length	PoF in stage of				
	15 days	30 days	45 days	60 days	90 days
Initial (none)	39.60%	43.78%	44.35%	43.82%	44.73%
4 m	22.30%	22.40%	22.40%	22.40%	22.40%
8 m	2.70%	2.40%	2.30%	2.30%	2.30%
12 m	2.30%	2.20%	2.20%	2.20%	2.20%

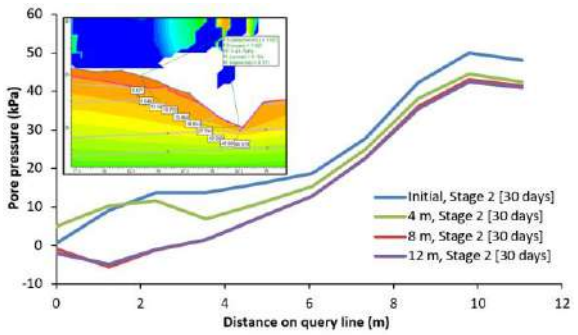


Figure 4. Pore pressure on each of the query points along query line for initial and various lenght of horizontal drain. Negative pore water pressures indicate groundwater level bellow the query points

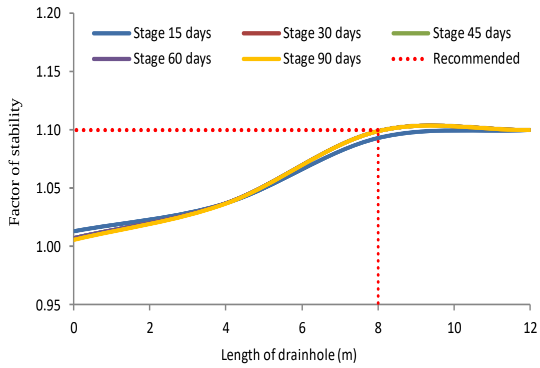


Figure 5. Corelation between length of horizontal drain and factor of stability on various stages of time

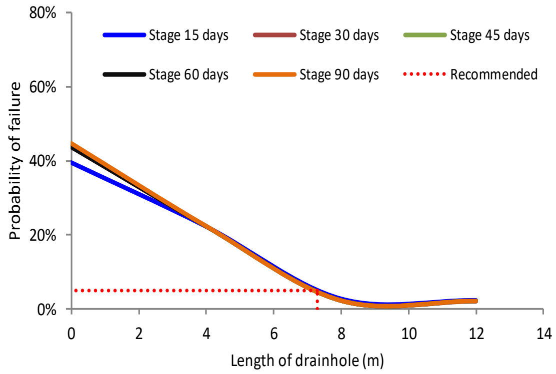


Figure 6. Corelation between length of horizontal drain and probability of failure on various stages of time

slope sections at different initial condition and horizontal drain. The query line was determined at fixed coordinate with 10 query points used to identify pore pressure at the contact zone.

Insert at the upper left of Figure 4 shows the query position in the analysis section for initial condition at 30 days stage. Based on the position sequentially from the top (see insert) or from the left (see chart), query points 1, 2, and 3 are query points for the top of the slip surface; 4, 5, and 6 for the middle; while 7, 8, 9, and 10 are for the bottom of the slip surface. A significant reduction in pore pressure occurred only at the installation of horizontal drain with the length of 8 m. After that, the changes in pore pressure were observed but it is not very significant.

To determine the effect of horizontal drain length on the values of stability factor and failure probability, modeling was carried out by correlating the stability factor (Figure 5) and the failure probability (Figure 6) with the horizontal drain length. In the correlation, the initial condition without drainhole installation is illustrated as the horizontal drain length of 0 m.

5. CONCLUSION

Installation of horizontal drain with specific designed perforated pipe can have a significant impact on slope stability. The groundwater from the formation can still come out to the surface or channel without carrying cement material, so the decline of sandstone cohesion does not occur. Then, with the discharge of water from the formation, the pore pressure decreases and the stability increases. The selection of horizontal drain length is one of the parameters that must be considered to produce optimal stability improvement. The longer the length, the better, of course; but it requires greater effort and cost. From the analysis, it can be seen that the horizontal drain with a length of 8 m and a slope of 3% is the optimal design to increase the stabilization. The horizontal drain may improve slope stability over time and this method is one of the optimum stabilization methods compared to other methods.

6. REFERENCES

1. Tao, Y. Z., Xi, D. Y. "Rule of Transient Phreatic Flow Subjected to Vertical and Horizontal Seepage". ***Applied Mathematics and Mechanics***, Vol. 27, No. 1, (2006), 59-65. DOI: 10.1007/s10483-006-0108-1
2. Price, J., "Implications of Groundwater Behaviour on the Geomechanics of Rock Slope Stability", in 1st Asia Pacific Slope Stability in Mining Conference, Australian Centre for Geomechanics, (2016), 25-48. https://doi.org/10.36487/ACG_rep/1604_0.3_Price
3. Reano, E., Beale, G., Dowling, J., Tejada, L. C., Lacey, M., Hazwezwe, H., "Development of A Mine Dewatering and Pit

Slope Depressurisation Review Process", in International Symposium on Slope Stability in Open Pit Mining and Civil Engineering 2020, Australian Centre for Geomechanics, (2020), 1253-1266. https://doi.org/10.36487/ACG_repo/2025_85

4. Rahardjo, H., Hritzuk, K. J., Leong, E. C. Rezaur, R. B. "Effectiveness of Horizontal Drains for Slope Stability". ***Engineering Geology***, Vol. 69, No. 3-4, (2003), 295-308. DOI: 10.1016/S0013-7952(02)00288-0
5. Hill, K., Hodgkinson, R., Harris, D., Newell-Price, P., Field drainage guide: principles, installations and maintenance. Warwickshire: Agriculture and Horticulture Development Board, 2018.
6. Hamedi, A., Marundi, S. M. "Laboratory Studies on the Effect of Vertical Gravel Column Drains on Liquefaction Potential". ***International Journal of Engineering, Transactions B: Applications***, Vol. 24, No. 3, (2011), 209-224. DOI: 10.5829/idosi.ije.2011.24.03b.02
7. Rahardjo, H., Santoso, V. A., Leong, E. C., Ng, Y. S., Hua, C. J. "Performance of Horizontal Drains in Residual Soil Slopes". ***Soils and Foundations***, Vol. 51, No. 3, (2011), 437-447. DOI: 10.3208/sandf.51.437
8. Campbell, R., Mackie, D., Anderson, W. S., "Integrated Slope Stability Assessment in A Complex Geotechnical and Hydrogeological Setting", in International Symposium on Slope Stability in Open Pit Mining and Civil Engineering 2013, Australian Centre for Geomechanics, (2013), 555-567. https://doi.org/10.36487/ACG_rep/1308_36_Campbell
9. Rahmannejad, R., Toufigh, V. "Influence of Curing Time and Water Content on Unconfined Compressive Strength of Sand Stabilized Using Epoxy Resin". ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 8, (2018), 1187-1195. DOI: 10.5829/ije.2018.31.08b.05
10. Mousavi, S. "Effect of Moisture Content on Shear Strength of Offshore Clay Interface Steel Surface". ***International Journal of Engineering, Transactions B: Applications***, Vol. 29, No. 2, (2016), 183-191. DOI: 10.5829/idosi.ije.2016.29.02b.07
11. Grimstad, G., Amiri, S. A. G., Nordal, S. "Relations and Links Between Soil Mechanics Porous Media Physics, Physiochemical Theroy and Effective Medium Theory". ***Frontiers in Physics***, Vol. 7, Article 41, 1-10. DOI: 10.3389/fphy.2019.00041
12. Rott, N. "Note on the History of the Reynolds Number". ***Annual Review of Fluid Mechanics***, Vol. 22, No. 1, (1990), 1-12. DOI: 10.1146/annurev.fl.22.010190.000245
13. Erskine, W. D. "Channel Response to Large-Scale River Training Works: Hunter River, Australia". ***Regulated Rivers: Research & Management***, Vol. 7, No. 3, (1992), 261-278. DOI: 10.1002/RRR.3450070305
14. Shao, X., Wang, H., Wang, Z. "Interbasin Transfer Projects and Their Implications: A China Case Study". ***International Journal of River Basin Management***, Vol. 1, No. 1, (2003), 5-14. DOI: 10.1080/15715124.2003.9635187
15. Yevjevich, V. "Water Diversions and Interbasin Transfers". ***Water International***, Vol. 26, No. 3, (2001), 342-348. DOI: 10.1080/02508060108686926
16. McEwan, A. L., "The Failure of and Remedials to A River Diversion for An Opencast Mine In The Witbank Coalfields of South Africa", in International Mine Water Association Congress 1999, International Mine Water Association, (1999), 79-85. Retrieved January 30, 2021, from http://www.imwa.info/docs/imwa_1999/IMWA1999_McEwan_079.pdf
17. Riyadi, F. A., Cahyadi, T. A., Nurkhamim, Supandi. "Desain Saluran Terbuka Berbasis Microsoft Excel Perhitungan dan Pemodelan yang Praktis dan Efisien". ***Kurvatek***, Vol. 4, No. 2, (2019), 61-78. DOI: 10.33579/krvtk.v4i2.1563
18. Flatley, A., Rutherfurd, I. D., Hardie, R. "River Channel Relocation: Problems and Prospects". ***Water,*** Vol. 10, No. 10, 1360, (2018), 1-25. DOI: 10.3390/w10101360
19. Supandi, Zakaria, Z., Sukiyah, E., Sudradjat, A. "The Correlation of Exposure Time and Claystone Properties at The Warukin Formation Indonesia". ***International Journal of Geomate***, Vol. 15, No. 52, (2018), 160-167. DOI: 10.21660/2018.52.68175
20. Supandi, Zakaria, Z., Sukiyah, E., Sudradjat, A. "The Influence of Kaolinite-Illite toward Mechanical Properties of Claystone". ***Open Geosciences***, Vo.11, No. 1, (2019), 440-446. DOI: 10.1515/geo-2019-0035
21. Supandi, Hartono, H. G. "Geomechanic Properties and Provenance Analysis of Quartz Sandstone from The Warukin Formation". ***International Journal of Geomate***, Vol. 18, No. 66, (2020), 140-149. DOI: 10.21660/2020.66.50081
22. Supandi, Hidayat, H., "The Impact of Geometry Bedding toward Slope Stability in Coal Mining", in 4th International Symposium on Geotechnical Safety Risk 2013, Geotechnical Safety Network, (2013), 559-562. https://doi.org/10.1201/b16058-85.
23. Supandi, Zakaria, Z., Sukiyah, E., Sudradjat, A. "New Constants of Fracture Angle on Quartz Sandstone". ***International Journal on Advanced Science, Engineering and Information Technology***, Vol. 10, No. 4, (2020). DOI: 10.18517/ijaseit.10.4.8272
24. Supandi, "Determination Material Properties on Bedding Contact at The Low-Wall Part of Coal Mine", in ISRM European Rock Mechanics Symposium 2014, ISRM, (2014), 903-907. https://doi.org/10.1201/b16955-155

Persian Abstract

چکیده

حفظ ثبات تأسیسات عمومی ، به ویژه جاده های عمومی ، بسیار ضروری است ، بنابراین تلاش هایی برای تثبیت باید انجام شود. آب حاصل از تراوش که از لایه ماسه سنگ خارج می شود ، مواد را با ذرات ریز حمل می کند و این باعث می شود انسجام سنگ کاهش یابد و در نتیجه منجر به شستشوبستر جاده شود. در نتیجه تمیز کردن ، لایه بیش از حد می شود و می تواند ثبات جاده را مختل کند. این مطالعه برای ارائه توصیه ای ساده اما قابل اندازه گیری برای حفظ ثبات جاده پس از شکست تلاشهای تثبیت قبلی انجام شده است. در این روش از روش مکانیک سیالات استفاده شده است که در آن به آب حاصل از فضای خالی شکل گرفته است بدون اینکه ذرات ریز (سیمان) از فضای بستر جاده تخلیه کند خارج شود. تجزیه و تحلیل نتایج نشان داده است که خروج آب با استفاده از روش اجزای محدود با نصب لوله تخلیه افقی انجام گردید. نتیجه تجزیه و تحلیل نشان می دهد که نصب زهکشی افقی به خارج شدن آب بدون ایجاد یک لایه بیش از حد کمک می کند. تخلیه افقی توصیه شده ۸ متر طول با شیب حداقل ۳٪ و طول ۱ متر است که بایستی منبع آب شکل گرفته از شن و ماسه و سنگ عبور کند و با استفاده از این روش می توان ثبات جاده را حفظ کرد و تثبیت بستر جاده را به روشی آسان ، ارزان و قابل اجرا در آورد.

IJE TRANSACTIONS A: Basics Vol. 34, No. 4, (April 2021) 1052-1065

International Journal of Engineering

Journal Homepage: www.ije.ir

Exploration of Rhenium Volcanogenic Deposit and Technology Development

Rylnikova M. V.[a], Argimbaev K. R.*[b], Ligotsky D. N.[b]

[a] *Research Institute of Comprehensive Exploitation of Mineral Resources of the Russian Academy of Sciences, Moscow, Russia Federation*
[b] *Faculty of Mining, Saint Petersburg Mining University, St. Petersburg, Russia Federation*

PAPER INFO

Paper history:
Received 30 November 2020
Received in revised form 30 December 2020
Accepted 15 January 2021

Keywords:
Rhenium
Volcanogenic Deposit
Scoop Conveyor
Geothermal Fields

ABSTRACT

Rhenium is widely used in manufacturing industry and metallurgy. Today the consumption of rhenium is high, but there are very few deposits in the world where it is mined. Thus, the aim of the study is to identify areas of distribution of rhenium on Kudryavy volcano, located on the islands of the Kuril ridge (Russia). In this connection, during the field period, we took samples of a volcanic massif weighing 70 kg, and also studied the geothermal fields with a pyrometer. Laboratory research included the study of composition of samples by the method of inductively coupled plasma mass spectrometry and spectrometric analysis. The article defines the zones of distribution of rhenium mineralization, presents the results of measurements of the temperature of geothermal fields and the elemental composition of technological samples. The geology, development technology is described, the analysis of the destruction of rocks, determined by the acoustic method, the specific resistance of breakdown during electrothermal loosening is given. The parameters of the tubular shovel conveyor at which the productivity of 88 m^3/h is achieved. The results obtained will make it possible to identify promising geothermal fields, determine the development technology, and contribute to the study of volcanic deposits.

***doi**: 10.5829/ije.2021.34.04a.32*

1. INTRODUCTION

By 2019, rhenium production outside the Russian Federation increased by 9 % compared with 2013. This happened mainly due to a significant increase in the output of rhenium in the USA, which is the largest producer and consumer of this rare metal in the world. In 2019 the United States produced approximately 72 % of all rhenium. From 2013 to 2019, the production in this country almost doubled and amounted to 23120 kg per year, and consumption reached 25140 kg per year (Figure 1) [1, 2]. The demand for this metal in the world is very high with limited mineral resources [3-13].

Taking into account the plans of some companies to expand their capacity, it can be assumed that if there is such demand for this metal, the production of rhenium in salts at US steel mills may increase by about 3–4 tons per year in the coming years [12-23].

*Corresponding Author Institutional Email: *diamond-arg@mail.ru* (Argimbaev K. R.)

There are few rhenium deposits in the world that can meet the demand for this metal. One of these deposits is the Kudryavy volcano, located in Russia. This is a volcanic type deposit that can give the world the valuable metal rhenium.

Volcanic deposits can serve as a source of metals such as Zn, Cu, Au, Mo, Ni, Co, Ir, Mn, Re [1-6].

A review of studies of volcanogenic deposits carried out by the world's leading scientists proved that they can be a source of minerals. These studies give a definition to such deposits, provide a diagram of a modern sulfide deposit located in the Pacific Ocean, a classification of metals, their geographical distribution, an assessment of the amount of metals, examples of large-tonnage volcanogenic massifs of sulfide deposits in the world, and describe the principle of the formation of such deposits [1-23]. Despite this, the distribution of rhenium in the geothermal fields of volcanic deposits has not yet been studied, moreover technologies for their development have not been presented.

Further research in this area related to the study of volcanogenic deposits is a solution to a major problem

Please cite this article as: Rylnikova M. V., Argimbaev K. R., Ligotsky D. N., Exploration of Rhenium Volcanogenic Deposit and Technology Development, International Journal of Engineering, Transactions A: Basics Vol. 34, No. 04, (2021) 1052-1065

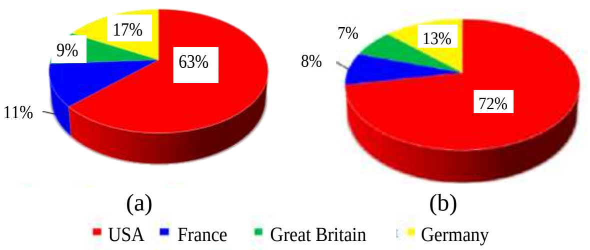


Figure 1. The share of foreign countries in rhenium production a. 2013; b. 2019

associated with an increase in efficiency of using mineral resources.

In this regard, this study is a relevant and timely direction of development for the purposes of providing rare elements to various industries.

The novelity of this research work is to establish the identification of patterns of rare elements accumulation in rocks located in high temperature zones, the dependence of the coefficient of rare element accumulation on the geothermal fields.

The limits of volcano crater rocks destruction by the seismic wave velocity were determined, which shows the possibility of using mechanical loosening. The dependence of the actual amount of material on the bucket capacity of excavation and loading equipment are defined at bucket fill factor *Kn*=0.4, 0.5, 0.6, as well as tubular scoop conveyor main parameters on the productivity of the mining complex and the density of the mineral was determined, which makes it possible to increase the efficiency of the complex and the profitability of the development.

The objective of the research is to identify areas of distribution of rhenium on Kudryavy volcano, as well as to describe the technology development and efficient transportation of rocks in difficult climatic conditions and at high temperatures.

The main results of the study passed semi-industrial tests and proved to be effective.

2. MATERIALS AND METHODS

The following tasks were solved in the course of research:

-Study of the material composition of rocks.

-Identification of geothermal fields with the highest temperature.

-Determination of main parameters of the technology for volcanogenic deposit development and rocks transportation.

The tasks were solved in two stages: field and laboratory.

During the field period, volcanic massif samples 70 kg each were taken from the geothermal site comprising Rhenievoye field, Angidridovoye field, Treschina field, Field 605, and Kupol field for further laboratory and technological study. The sampling place was a gas flow outlet on the flattened surface of a domal up warping in the Kudryavy Volcano vent.

The depth of sampling using bulk method varied from the daylight surface up to 0.5 m. The transfer to the «hot» zone of the gas jet discharge explained the sampling depth taken with a temperature exceeding 700˚C, which was measured using an external thermocouple. Samples were taken from a conical excavation with wall cutback up to 10 meters and narrowing to the depth. The sample comprised all observed rock differences without selecting preferable ones. The sampling area length was up to 400 meters in each geothermal field; samples were taken in 20 meters on average (Figure 2).

Temperature parameters of geothermal fields were measured using AKIP-9307 manual high-temperature pyrometer with measuring values up to plus 1000˚C. Measurements were made under minimal fumarole activity on the open surface site. Measurement results were recorded on the topographic surface.

Topographic survey of the vent relief was scaled 1:1000 with relief sectioning by horizontal intervals in each 1 meter using tacheometric procedure from survey control points by Sokia SET 630R electronic tacheometer. Topographic survey accuracy was checked by inspecting field data and comparing sketches and photographs. Field data was verified in terms of registration and completion of logs. The operating procedure was checked as well as tolerances for survey justification and the topographic survey procedure.

Then, laboratory tests were carried out for the samples taken, which were both coarse aggregates sizing more than 100 mm and fine-dispersed component (ash).

Taking into account the non-uniform grain-size distribution of samples, a set of preparative procedures was carried out comprising multistage crushing of the

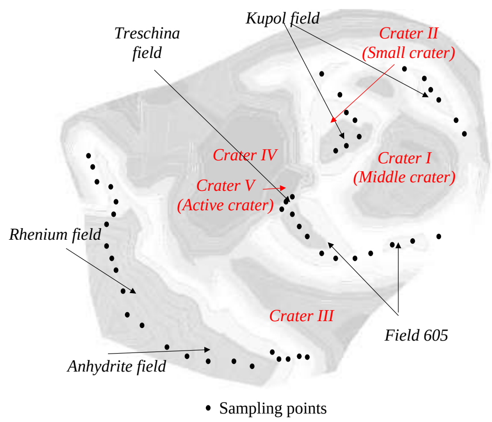


Figure 2. Sampling layout

initial material to fineness -1 mm with preliminary and control sieving. Figure 3 illustrates the preparation flowchart for the samples to be studied.

As a result, samples with uniform grain-size distribution were prepared.

An element composition of process samples was studied using mass spectrometry with inductively coupled plasma (ICP-MS) and NSAM (Analytical Method Scientific Council) industry procedure No. 179-X "Rhenium Measurement in Rocks and Sulphide Ores by Photometric Method (2015), Russian Federation."

The following operating conditions applied for Agilent-7500 ICP-MS: plasma temperature 8000 to 10000˚C; carrier gas flow rate 0.8 to 1.3 L/min; plasma-supporting gas flow rate 15 L/min; high-frequency signal power 700 to 1600 W; signal integration time 0.1 s.

To calibrate the mass spectrometer and perform the assay, laser ablation was carried out for two zones in the sample. The results obtained in such a way were compared with certified element contents (Ag, As, Au, Br, Ba, Be, Bi, Ca, Cd, Ce, Cl, Co, Cr, Cu, Dy, Er, Eu, Fe, Ga, Gd, Ge, Hf, Hg, Ho, I, La, Li, Lu, Mg, Mn, Mo,

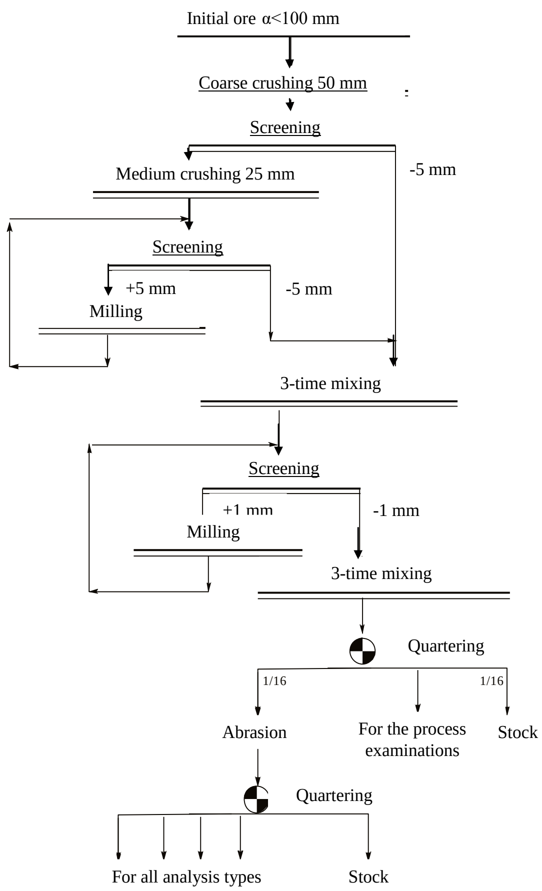


Figure 3. Preparation flowchart for the samples to be studied

Nb, Nd, Ni, Pd, Pb, Pr, Pt, Ru, Sb, Sc, Sm, Sn, Sr, Ta, Th, Ti, Tl, U, V, W, Zn, Zr). There were no systematic deviations; data variability did not exceed procedural error limits.

The samples featured high contents of molybdenite, titanium, iron and manganese that distorted the rhenium values measured using ICP-MS. Matrix effect affected the procedure accuracy. To minimize the matrix impact, it is necessary to dilute the samples in high proportion, but it was impossible to avoid the molybdenite impact completely. Rhenium separation from the matrix by isotope dilution is the most common procedure to minimize the matrix impact. Isotope dilution has several limitations: first of all, the element to be measured should have at least two isotopes; secondly, those isotopes should be free from spectral interferences. Despite the progress in ICP-MS technology over the last decades, high costs of equipment limit this procedure for routine assays.

To measure the rhenium contents, we used NSAM No. №179-X, which is based on spectrophotometric analysis and is relatively affordable, accessible and reliable ensuring the measurement of rhenium concentration within 10^{-2} % wt. to 10^{-5} wt% [24,25,26]. The applied procedure is standardized and certified. This method is fast and easy when executing [27]. It comprises the preparation of a rhenium (VII) solution, reducing it to rhenium (IV) with a tin (II) chloride solution, them converting it into a complex compound, sorbent separation from the solution, diffuse reflectance measurement at 510 nm, and calculation of rhenium (IV) contents using a calibration curve. Silica modified chemically with N-(1,3,4-thiodiazole -2-thiol)-N'-propylurea groups was used as a sorbent.

To establish calibration parameters of rhenium measurement, a reference solution (state standard reference sample) No. 100043-2 was used. Solutions with lower concentrations were prepared by dissolving aliquots of the reference solution. The used procedure had the following metrological characteristics specified in Table 1.

Quality indicators of NSAM No. 179-X procedure (Table 1) satisfied metrological requirements, and the results obtained were reliable.

Application and efficiency of the volcanic massif mechanical loosening were examined by seismoacoustic method based on the study of propagation of elastic vibrations in the massif.

The speed of elastic wave propagation sufficiently correlates with the massif strength and jointing and may serve as a generalized indicator that considers the change of these factors.

Elastic wave propagation speed increases with the rock strength increase and decreases with jointing increase.

TABLE 1. Quality indicators of NSAM No. 179-X rhenium measurement procedure (weight fraction, %)

Rhenium content, % wt.	Repeatability		Intra-laboratory precision		±Δ	Reproducibility		Uncertainty value		
	S_r	r	$S_{I(TO)}$	$R_{I(TO)}$		S_R	R	N	U_c, % wt.	U
0.05	0.0009	0.0025	0.0011	0.0030	0.0018	0.0014	0.0040	0,0140	0,007	0,014
0.10	0.0012	0.0033	0.0014	0.0040	0.0024	0.0019	0.0050	0,0323	0,0034	0,0068
0.50	0.0060	0.0170	0.0070	0.0200	0.0120	0.0110	0.0300	0,0600	0,0034	0,0068

Note: S_r, $S_{I(TO)}$, S_R are standard deviation of repeatability, intra-laboratory precision and reproducibility; r, $R_{I(TO)}$, R are repeatability, precision and reproducibility limits; ±Δ are the limits of the range for the analytical result error at confidence factor P=0.95, N is relative standard uncertainty, U_c is total standard uncertainty, U is extended uncertainty of analytical results.

To exclude the impact caused by the properties of the rocks forming the massif on the seismic wave propagation speed, the ratio was used between velocities of longitudinal waves with the same type defined by measuring at large 30 m bases and at bases approx. 10 cm in monolithic areas of the same massif. Measurements were carried out using a high-frequency mic survey in the first case and by ultrasonic method in the second one. The velocities obtained were designated as $V_{\text{high frequency}} = V_{\text{ultrasonic}}$, correspondingly; their ratio will define the acoustic jointing indicator, i.e. $R = V_{\text{high frequency}} / V_{\text{ultrasonic}}$. $V_{\text{high frequency}}$ may serve as an indicator taking into account the structural characteristics of the rocks to be loosened. Assessment of the massif by two parameters, R and $V_{\text{ultrasonic}}$, considers the most comprehensively the interaction between the ability to loosen and physical-technical parameters of the massif.

Laboratory research included the study of composition of samples, technological tests of new equipment for the transportation of rocks at various angles of inclination from 0° to 180°, according to the description of Russian patent RU No. 170400, as well as an electrothermal method of selected samples destruction using aluminum cathodes, IOM 100/25 high-voltage transformer (single-phase oil test transformer with a frequency of 50 Hz) with a rated power of 25 kVA, and IOM 100/100 transformer with a rated power of 100 kVA.

3. RESULTS

3. 1. Geological Structure of Kudryavy Volcano

Kudryavy Volcano is 986 m high and its base diameter is about 5 km. It is superimposed on an ancient, heavily reworked volcano structure. Including the Sredney volcano. All rocks of the volcano are subdivided into three strata of different ages. The lower thickness consists of numerous blocky lava flows of andesite-basalt and andesite composition, exposed at the periphery of the volcano and separated from each other by breccia crusts. The extruded sheet occurred mainly in the western direction; to the east, the spread of lavas was limited by the buildings of the rhyolite dome and Sredney volcano. The surface of the lava flows is overgrown with elfin, alder, birch, especially at elevations less than 350 m. After the formation of this stratum, the Kudryavy Volcano, apparently, was in a state of relative rest. Slope breccias with interlayers of tuffs and organic matter formed on the southern and northern slopes.

The section of the sediments overlying the lavas of the lower strata described by us is in a dry ravine between the extrusive body and the lava flow of the last eruption on the northern slope of the Kudryavy Volcano.

On the lavas of the southern slope of Kudryavy volcano, there is a lens composed of interbedded loam with peat bogs. In total, there are 5 interlayers of peat bogs with the thickness of 3-4 cm. The age of the second peat bog on top is 4450 ± 100 years. Thus, the age of the lower complex of Kudryavy Volcano is possibly over 5000 years old.

The middle stratum is composed of highly oxidized viscous lavas and an andesitic extrusive dome. The distribution of this stratum is limited by the near-crater part of the volcano. It is described in more detail in the description of the peak Kudryavy Volcano.

The upper thickness includes 7 flows of olivine-bearing bipyroxene basaltic andesite and andesite, traced towards the northwest, west and south of the peak. Its lavas are separated from the middle strata by a section that includes lacustrine-swamp deposits and pumice. It is the most fully sketched in (from top to bottom) on the northern slope of the Menshoy Brat Volcano. It follows that the marking pumice horizon has an age of <700 ± 40 years, and the upper age limit of the sequence is limited by the dating of uncarbonized elfin wood (170 ± 40 years), underlying the largest flow of block lavas of the upper sequence.

At the top of the Kudryay Volcano, according to geological and morphological features, elements of four craters and an explosion funnel are distinguished, they are differing in age, structure and mode of fumarolic activity (Figure 2).

Crater I is a middle crater (230 m in diameter) and the most ancient (it forms the lower stratum), is represented by a somma fragment in the center of the somma of Kudryay volcano. Spatially it limits all high-temperature sites of the Kudryay Volcano. Its somma is composed of interbedded basaltic and andesite lavas and agglomerates strongly altered by solfatary-fumarolic activity. The rest of the crater was destroyed by later eruptions.

Crater II is superimposed on the first one and is located in the eastern part of the volcano. In its outlines, it is close to a regular circle (130 m in diameter along the upper edge of the somma), it is open to the east. The somma is formed by viscous andesite lavas (middle stratum) with inclusions and veins of cordierite glass. Andesite extrusion (Kupol field) is located inside the crater, occupying more than half of the crater area. It is associated with 3 fumarole sites with high-temperature gases, which received their own names (in brackets: area in square m; average / maximum t ° C): «Treshina field» (322; 528/750), «Kupol field» (980; 620/920) and «Field 605» (396; 586/784) - confined to the upper edge of the somma

The lowered parts of the relief are filled with crater-lacustrine deposits, the formation of which is due to erosion (to a significant extent is the wind) of the surrounding hills and temporary flows associated with melting snow and heavy rainfall. The depression to the west of the extrusion dome is quite hot (52 ºC at a depth of 5 cm). With a deepening of 1.5 m, the temperature rises up to 96 ºC and is accompanied by intense release of sulfurous gases. Possibly, sandy-argillaceous deposits are the barrier to gas jets and form the cover that prevents the dispersion of volcanic gases and directs their movement to the «Rhenium field site». In other depressions during pits driving (up to 2 m), the temperature rise was not recorded. These depressions are likely traces of large explosion craters that formed around the Kupol field, similar to the crater in October 1999, but later filled with tephra and sediments of temporary streams.

Crater III is 80 m in diameter. It is located to the west of the first one and is superimposed on the somma of Crater I of the Kudryavy volcano. Its formation is associated with weak eruptive activity and the intrusion of a small stock-like body of andesites, now completely processed by high-temperature gas jets. The somma of Crater III is also practically not preserved, and its position is reconstructed fragmentarily. In its western part there are two fumarole sites (in brackets: area in m^2; average / maximum t ºC): «Rhenium field» (1048; 481/020); «Anhydrite field» (140; 280/360).

Crater IV spatially limits the distribution of low-temperature fumaroles and has the shape of an ellipsewhich is open to the north. The last three lava flows erupted from it towards the northern, southern and southwestern directions (upper strata), and andesite-basaltic slags were thrown out, overlapping all formations of the volcano. Curly, several isolated areas with fumaroles are also distinguished here, the gases of which have temperatures predominantly of 98-120 ºC (up to 21.0 ºC) and intensively deposit sulfur. However, in some areas, the gas temperature exceeds the melting point of sulfur.

Crater V was formed as a result of the eruption on October 7-8, 1999, which began as a phreatic [30, 31]. Later, Znamensky et al. [28] informed about lava observed at the bottom of the crater in the second half of October. Crater V is an explosion funnel with a depth of 15 to 40 m, an upper circumference is about 15 m in diameter and almost sheer walls. The funnel cut off a part of the Crater II ridge and exposed a part of the exocontact of the Kupol field extrusion. At its bottom, an eruption of gases from a hole in red-hot (to a red glow) rocks was observed. Although all high-temperature fumarole sites are located within a radius of 200 m from this funnel, its explosion did not affect the mode of their degassing to a visually noticeable degree, while the extinct fumarole of the dacite dome was activated.

3. 2. Geothermal Fields of Kudryavy Volcano and Rare Metal Mineralization Scientists' research proves the accumulation of minerals in areas of high temperatures of the volcanogenic massif [32,33,34]. Such a phenomenon was identified in the geothermal fields Kudryavy volcano too.

The study of geothermal fields was carried out along all profiles where samples were taken, which made it possible to identify the highest temperature of 900 ºC that was observed in the Kupol field.

The results of measurements are presented in the form of a diagram of volcano crater surface temperature distribution in Figure 4.

Fieldwork and analytical works showed that rare metal mineralization forms in high-temperature geothermal fields and is distributed over a total area of 7,600 m^2, which can exceed bulk earth values. The low-temperature region of the volcano is distinguished by little rare-metal mineralization. The nature and composition of accumulations are directly dependent on the temperature and structure of adjacent rocks. The most favorable rocks for such accumulations are porous slags, as well as fractured rocks and tectonic breccias, which attract and deposit minerals, which is confirmed by visual inspection during sampling.

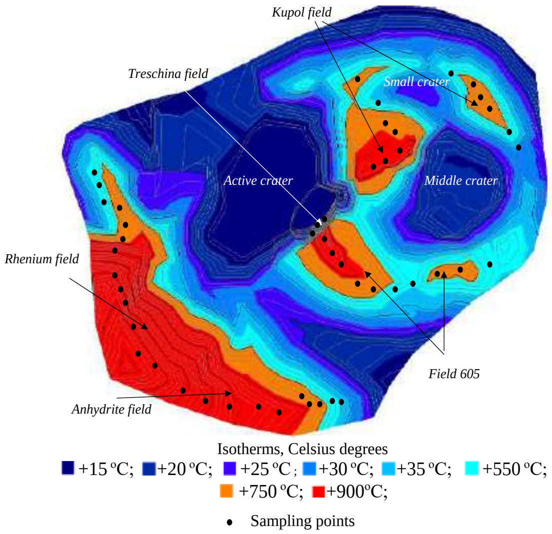


Figure 4. Diagrams of Kudryavy volcano crater surface temperature distribution (Russia)

The results of the analysis of the elemental composition of technological samples by the ICP-MS method (Table 2) showed that the samples are characterized by an increased content of molybdenum, titanium, iron and manganese, and a lower content of tungsten and zinc. Each sample was analyzed from two portions - quartered from abraded samples of 2 kg (6112 and 6115) and samples of 50 g taken by the standard method (6112-1 and 6115-1). Comparison of the obtained data confirms the uneven distribution of valuable components and the need for more thorough mixing of the material in the process of analytical research.

According to the results of the analysis of initial samples by method NSAM No. 179-X, the content of Re in them is as follows: in sample 6112 - 280.5-387.7 gpt, in sample 6115 - 10.43-25.72 gpt. (Table 3).

TABLE 2 Results of analysis of the elemental composition of technological samples by ICP-MS

Element	**Content, ppm**			
	Sample No.			
	6112	**6112-1**	**6115**	**6115-1**
Be	1.75	0.775	1.83	1.04
Ti	2366	3437	3721	3984
V	300	257	292	232
Cr	32.5	33.8	23.1	29.5
Mn	1840	1347	1341	1384
Fe	142134	63258	79633	79490
Co	32.0	15.8	21.8	25.4
Ni	25.2	14.4	15.1	20.2
Cu	308	296	549	171
Zn	760	80.3	1047	4.14
Ga	20.7	22.2	21.2	17.6
Rb	52.6	6.56	26.4	15.2
Sr	272	230	308	229
Y	22.7	21.9	23.1	20.1
Zr	44.3	53.6	53.7	49.9
Nb	0.988	0.642	0.916	0.694
Mo	15973	125	3299	206
Sn	49.6	3.83	9.47	2.12
Cs	3.93	0.155	1.56	0.189
Ba	93.9	122	206	150
La	3.39	3.96	4.20	3.86
Ce	9.76	10.7	11.5	10.6
Pr	1.56	1.67	1.81	1.61
Nd	8.25	7.78	9.02	7.74
Sm	2.64	2.71	2.72	2.36
Eu	0.684	0.621	0.823	0.759
Gd	2.86	2.94	3.20	3.03
Tb	0.590	0.570	0.626	0.553
Dy	3.92	3.68	3.97	3.45
Ho	0.885	0.872	0.889	0.793
Er	2.73	2.60	2.63	2.38
Tm	0.424	0.422	0.400	0.383
Yb	2.88	2.73	2.87	2.32
Lu	0.470	0.408	0.431	0.371
Hf	1.30	1.41	1.51	1.41
Ta	0.228	0.075	0.083	0.094
W	525	5.18	242	6.80
Th	0.605	0.598	0.515	0.756
U	0.510	0.359	0.411	0.338

Note: There were no systematic deviations; data variability did not exceed the procedural error limits.

TABLE 3 Determination of the content of Re in the initial samples by method NSAM No. 179-X

Sample No.	**Re content, gpt**	S_r	**±Δ**	**U**
6112	387.7	0.0010	0.0018	0.012
6112-1	280.5	0.0013	0.0023	0.0073
6115	25.72	0.0008	0.0016	0.0074
6115-1	10.43	0.0015	0.0025	0.0068

Some of the studied elements can be differentiated by geothermal field depending on the temperature (Figure 5).

Thus, it can be seen from the Figure 5 that the maximum Re concentrations (up to 300 g per ton) were found in the Rhenium field ores, and formed at 560 ºC. Three peaks are distinguished in the Re distribution. The main one (at 500-620 ºC) is provided with samples from the «Rhenium field», where rhenium disulfide itself makes the greatest contribution. Somewhat less (at 620-720 ºC) is due to samples from Crater II. Here Re is positively correlated with Zn (0.50), Cu (0.37), W (0.26), Ge and Mo (0.19). Rhenium disulfide was not detected in this temperature range, but there are single determinations of rhenium in molybdenite and sphalerite, and in the sublimates of quartz tubes, K-Re oxide was found, which is close in composition to $3KReO_4*ReO_3$. Although complete isomorphism in the Mo - Re series is assumed for the former, the absence of a significant correlation between these elements forces us to assume that sphalerite is also a Re concentrator in high-temperature ores. The third maximum stands out in samples at <400 ºC and is associated with Re impurities in Pb-Bi sulfosalts and pyrites.

Increasing concentrations of In (>1000 g per ton) is typical for samples deposited at 450-600 ºC, mainly at «Rhenium field». Positive correlations of In with Cd (0.70), Zn and Sn (>0.45) indicate the predominant deposition of native In minerals (sulfides and sulfo salts) in this temperature range, with the weak deposition of In halides in the low temperature range.

Mo-bearing mineralization occurs in two areas of deposition. The first bearing, with Mo contents in most samples >1000 g per ton, represents the high-temperature areas of the Crater, and the second one, with Mo concentrations <800 g per ton, is represented mainly by samples from the Rhenium field. In both cases, the direct dependence of the Mo content on temperature is clearly manifested. High concentrations of molybdenum in the ores of Crater II are due to massive crystallization of molybdenite (at temperatures >620 ºC) and other Mo-containing minerals (ilsemanite, novellite, molybdite, etc.) at lower temperatures.

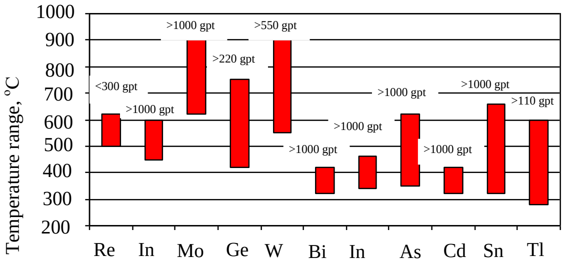


Figure 5. Temperature range of elements accumulation in rocks

The maximum Ge concentrations (100-220 g per ton) are characteristic of the Field 605 mineralization formed in the temperature range 420-750 ºC. Minerals of Ge have not yet been found, but not excluded. The high values of the correlation coefficients of Ge (at the critical value $r_{0.05}$ = 0.2732 for a selection of 51 samples) with Cu (0.68), Zn (0.60), Ag (0.38) are explained by its occurrence in sphalerites and sulfosalts, while with Mo (0.62) and W (0.45) are allowed to suggest an isomorphic incorporation of Ge into the minerals of tungsten and molybdenum. Although the latter are one of the important carriers of Re in ores, the correlation coefficient of Re and Ge is estimated to be insignificant (-0.02 at $r_{0.05}$= 0.2732) for volcano. Curly in general. However, in the most renaissance sites «Field 605» and «Rhenium field», the correlation of Re and Ge is estimated by the coefficients + 0.19 (at $r_{0.05}$ = 0.1954 for the sample of 69 samples) and + 0.26 (at $r_{0.05}$ = 0,1638 for the sample of 147 samples). Diagrams of correlations in ores Kudryavy volcono ones show the incorporation of Re into the Re-Ag-Cd-InCo-Ni association, and Ge - into the Zn-Ag-Ge-Ge-Ga-Cu-Mo-W association. These associations form mineralization in different temperature zones, but are linked through Ag. The fact of joint accumulation of Re and Ge in ores (at Re/Ge = 1.9) attracts attention as an unusual phenomenon, previously noted only in the conditions of the bituminous metalliferous formations of Mansfeld and Colorado [32]. It must be of special consideration in the conditions of the South Okhotsk region, where large deposits of germanium are known, and now Re-bearing objects have also been identified.

Concentrations of W (> 500 g per ton) are inherent in ores deposited at temperatures of > 550 ºC on all high-temperature areas. The intrinsic minerals W are represented by scheelite, povellite, stolcite, and tungstenite [34]; Hubnerite (Fe, Mn) WO_4 also sublimes in the tubes [34]. The positive correlation of W with Ge (0.45), Mo (0.26) and Re (0.19) indicates similar physicochemical conditions for the crystallization of minerals-carriers of these elements

In general, rare metal ores of high-temperature sites are characterized by low Cu concentrations, with the Zn/Cu ratio being one of the highest for the ores of the Kuril Islands. The maximum concentrations of Cu at high-temperature sites are inherent in the ores of Crater II, deposited at temperatures >550 ºC. No native Cu minerals were found here, and preparations of metallic copper «dissolve» in steam-gas jets with the removal of copper in the form of chlorides. The positive correlation of Cu with In, Ge, Sn, Re (0.30) and with Mo (0.45) indicates the possibility of the joint occurrence of these elements in the pyrites and pyrrhotites deposited at these temperatures. At the same time, in the sulfur-sulfide sublimates of Crater IV, copper forms its own minerals (chalcopyrite, covellite, bornite, chalcocite) or impurities in Fe sulfides (pyrrhotite, pyrite, marcasite).

The maximum concentrations of Zn (>1000 g per ton) are typical for ores of all high-temperature areas and are provided by the development of zinc sulfides (sphalerite, ZnCdIn-minerals, etc. [32]), less oxychlorides [34], There is a direct correlation between the Zn abundance and temperature.

The maximum concentrations of Pb (> 10 000 g per ton) are inherent in the Rhenium field ores. They are due to the development of intrinsic minerals, mainly Pb-Bi-Se sulfosalts, and in pipes, at relatively low temperatures, also chlorides [34], which is reflected by the negative correlation between Pb and temperature.

The maximum Bi concentrations (>1000 g per ton) are characteristic only for the Rhenium field and show an inverse dependence on temperature. Two peaks can be distinguished at 320-420 ºC and 500-650 C, indicating two temperature ranges and different modifications of the deposition of bismuth minerals: sulfide-sulfosaline and oxychloride.

The maximum As concentrations (>1000 g per ton) practically coincide with the Bi maximums in the Rhenium field ores deposited at temperatures of 340-460 ºC. Here, a positive correlation of As with Sb (0.40), T1 (0.38), Bi (0.21), indicating the probable dominance of Sb-As sulfosalts, the possible presence of lorandites and sulfates.

The maximum concentrations of Sb (30-100 g per ton) are also characteristic of the «Rhenium field» in the temperature range 350-620 ºC. Positive correlation coefficients of Sb with As, Bi, Sn, Ga (>0.20) and negative - with temperature evidence in favor of Sb deposition in low-temperature eulphosalts. Attention is drawn to the very high value of the As / Sb (s85) ratio in ores, probably exceeding As/Sb in most other volcanoes and ores of the Kuril islands.

The maximum Cd concentrations (>1000 g per ton) are typical mainly for the Rhenium field. As well as for Bi, two ranges of Cd concentration are clearly distinguished: at 320-420 ºC and at 500-650 ºC. This testifies to the presence of at least 2 mineral species - Cd concentrators. Of these, sphalerites, ZnCdIn sulfides, and greenockites are now known, and also chlorides in pipes [32].

The maximum concentrations of Sn (>1000 g per ton) are characteristic only for the Rhenium field in a wide temperature range (320-660 ºC). No intrinsic Sn minerals have been found, but both cassiterites and sulfostannates (at high temperatures) and chlorides (at lower temperatures) are assumed.

The content of Tl shows a good inverse dependence on temperature and maximum concentrations (> 110 g per ton) in the temperature range 280-600 ºC. The positive correlation of T1 with As (0.38) testifies in favor of the predominant occurrence of T1 in low-temperature lorandites together with As, Thallium chlorides were also found in the pipes. [32]

The general representation of the dependence of elements accumulation in rocks can be characterized by the accumulation coefficient normalized by the average composition in the earth's crust (Figure 6).

From Figure 6 it follows that in ores (Kn) decrease in the sequence: >1000 for Bi, Re, Cd, Mo, In; 100/1000 for W, Sn, Ag; 10/100 for Tl, B, Zn, Ge, Au, Sb, Cl, Pb, F, Cu.

The above figures characterize only the general tendencies of the accumulation of elements in the ores of the Kudrayvy Volcano. According to individual analyzes of certain mineral types of ores, Kn may be 1-3 orders of magnitude higher. Attention is drawn to the fact that the largest Kn have elements with the smallest clarkes and possessing chalcophilic and siderophilic properties.

3. 3. Technology of Development of Volcanic Deposits (Kudryavy Volcano)

On the basis of the studied development technologies [35-39], a new development technology with further transportation of minerals has been created. This technology includes layer-by-layer mechanical loosening, and when areas that are not amenable to mechanical loosening appear, the electrothermal method of destruction is used to create a system of cracks. The ripper is mounted on a bucket of earthmoving equipment that digs up minerals and loads them onto a conveyor (Figure 7).

Based on the results of field tests carried out on the geothermal massifs of rocks of the Kudryavy volcano, a diagram of manual jackhammer efficiency plotted according by the velocity of seismic waves were developed (Figure 8).

Figure 8 shows that geothermal volcanic rock mass has different strengths even in rocks of the same type. This is explained by the different rate of fumarole activity (the higher the activity, the higher the porosity of the material) and the temperature zone in the geothermal field. Therefore, a low seismic velocity

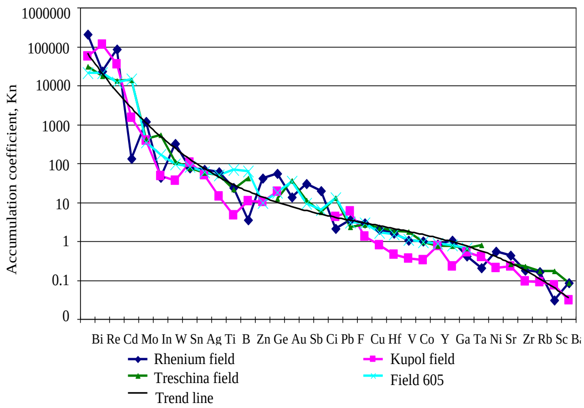
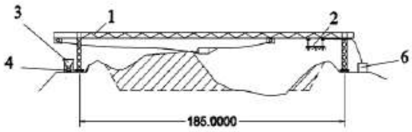
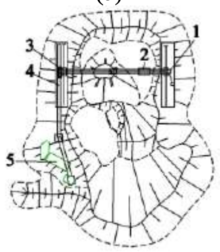


Figure 6. Coefficient of accumulation (Kn) of elements in the rocks of Kudryavy volcano



(a)



(b)

Figure 7. Principle scheme of Kudryavy volcano geothermal fields development using a modernized scraper a - transverse view; b - plan view; 1- modernized scraper with a suspended electrothermal installation; 2 - electrothermal installation; 3 - mobile bunker; 4 - conveyor; 5 - working platform with rock rehandling; 6 - power generator.

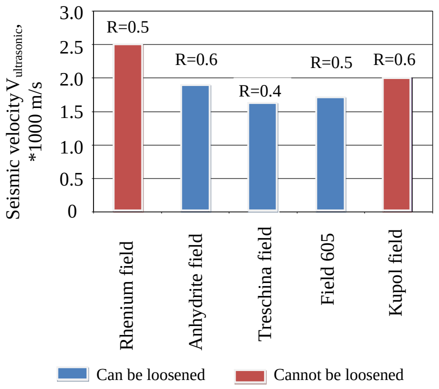


Figure 8. Diagram characterizing the destruction of rocks using an electric jackhammer, with a power of 1.75 kW and a mass of 16.5 kg, according to the speed of seismic waves

(up to Vultrasonic = 1800 m/s and an acoustic fracture index of R = 0.6) in ores of geothermal fields serves as an indicator of possible crushing using a mechanical ripper based on a ladle [44-47]. This crushing method may be ineffective in the Rhenium and Kupol fields due to solid inclusions.

These inclusions are well exposed to loosening by the electrothermal method [45–48]. It has been experimentally established that the specific resistance of breakdown in solid inclusions is 4.3-5.0 kOhm/cm when using a high-voltage transformer of the IOM 100/25 type (single-phase oil test transformer with a frequency of 50 Hz) with a rated power of 25 kVA, and with a transformer of the IOM 100/100 type, the specific breakdown resistance is 0.9 - 1 kOhm/cm.

As a result of a rupture, a system of cracks forms in the sample, which split it into several parts. If the sample is thick enough (more than 100 mm), then the general property of solid inclusions in which there is a breakdown is the complete regeneration of its dielectric properties after the voltage is removed. Subsequent exposure of the rock sample to the current does not cause the formation of a breakdown along the same trajectory, but is formed in a new region of the dielectric. This is due to the fact that as a result of melting and subsequent crystallization, a new substance is created, which has the more ordered structure and, as a result, lower conductivity. The breakdown channel in the crystalline state includes the air cavity surrounded by a tube of crystalline substance and the adjacent annealed part of the sample.

After loosening by a mechanical or electrothermal method, the rock is dug out using a bucket fixed with cables between two supports, which is driven by two winches. At the same time, an experiment carried out in laboratory conditions with selected samples revealed the filling factor of a ladle with a volume of up to 10 m^3 in the range of 40-60 % (Figure 9). Using of large buckets is not technologically easy.

Figure 9 shows that an increase in the lumpiness of extracted minerals can lead to a decrease in the productivity of mining equipment by reducing the actual amount of material in the bucket. This, in turn, leads to a decrease in productivity and all economic indicators.

To increase the productivity by minerals distinguished by increased lumpiness, it is necessary to perform additional loosening by the mechanical method. Then the maximum possible theoretical productivity of the mining complex can reach up to 1 million tons/year taking into account the actual amount of material in the bucket and the mode of occurrence.

The extracted material from the geothermal fields will be transported to the caldera, where the main production site with a tubular scoop conveyor is planned to be. It will allow you to overcome steep slopes,

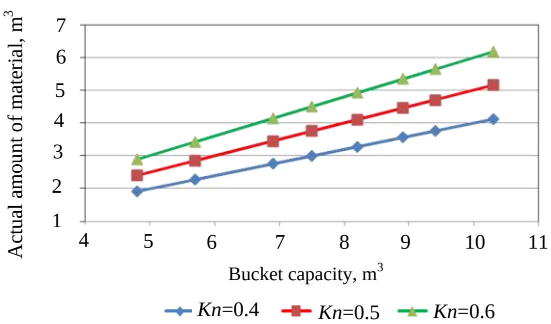


Figure 9. Dependence of the actual amount of material on the bucket capacity of excavation and loading equipment are defined at bucket fill factor Kn=0.4, 0.5, 0.6

preventing slipping, clogging, freezing and wetting of the transported material. Thanks to the cooling jacket located on the tubular body or the spiral heating element with thermal insulation coating, the device can operate in a wide temperature range from +50 to -60 ºC. The description of the new conveyor is given in the patent (RU No. 170400) (Figure 10).

At the same time, the main parameters of the conveyor (Figure 11) depend on the productivity of the mining complex and the density of the mineral, which was determined empirically and amounted to 7.5 t/m^3 for rhenium and from 4.0 to 7.0 t/m^3 for the rest of the minerals.

The analysis of the results obtained (Figure 11) allows to determine optimal parameters at which the minimum energy consumption and wear rate are observed.

Therefore, if the mineral density of 6.0 t/m^3 prevails, the transportation capacity can reach 88 m^3/h at a speed of 2.0 m/s, the scoop filling capacity of 80 % and the diameter of the outer tubular body of 160 mm

The calculated coefficients and actual loads carried by the scoops are derived from the condition of rocks supply to the conveyor using a feeder or other means that ensure the normal loading of vehicles. Such an organization of cargo flows is accepted for tubular scoop conveyors based on the calculations under similar conditions. When designing taking into account the specified productivity and drive power, the size of scoops and other parameters of the conveyor shall be chosen considering the calculated resistance to scoops motion caused by loading conditions, as well as the load that may arise at the transportation route.

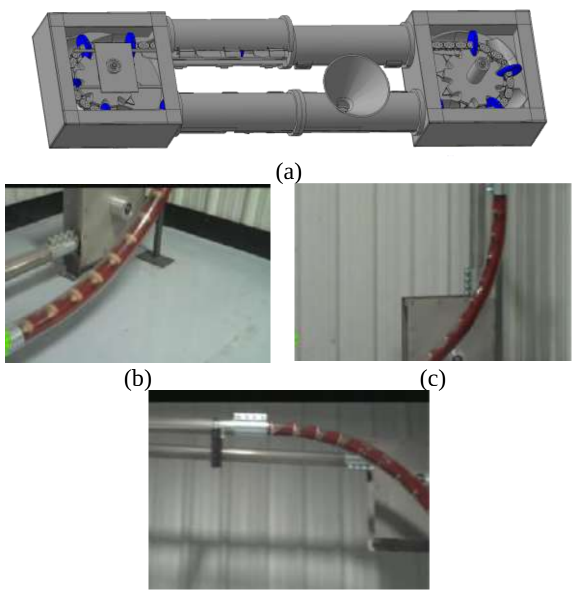


Figure 10. Tubular scoop conveyor (a) computer 3D model; (b) laboratory installation (transfer of the transported material from the horizontal section to the inclined one); (c) laboratory installation (the transported material overcomes the vertical section); (d) laboratory installation (transfer of the transported material from the vertical section to the horizontal one)

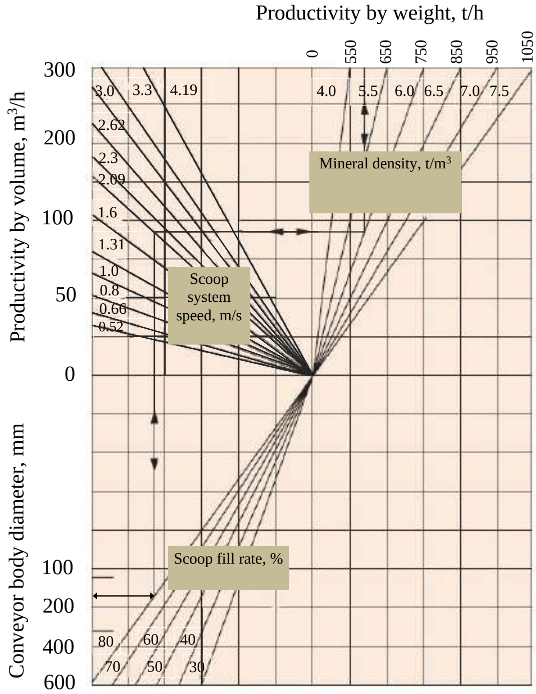


Figure 11. Dependence of the parameters of the tubular scoop conveyor on the productivity

4. DISCUSSION

Rhenium consumption increased by 9% in 2019 as compared with 2013. The USA is the major manufacturer of this metal with 72% of total production worldwide. At the same time, an increase in consumption will be about 3 to 4 tons annually over the coming years. Generally, this metal is very high in demand while its reserves are limited.

Rhenium volcanogenic deposit located in the Kudryavy Volcano vent (Kuril Islands) may become an additional source of this mineral, where areas of accumulation of large amounts of rhenium have been identified.

All rocks of the volcano are subdivided into three strata of different ages. The lower sequence consists of numerous blocky lava flows of andesite-basalt and andesite composition, exposed at the periphery of the volcano and separated from each other by breccia crusts. The outpouring of lava flows mainly in the western direction; to the east, the spread of lavas was limited by the buildings of the rhyolite dome and Sredny Volcano.

The middle stratum is composed of highly oxidized viscous lavas and an andesitic extrusive dome. The distribution of this stratum is limited by the near-crater part of the volcano.

The upper thickness includes 7 flows of olivine-bearing bipyroxene basaltic andesite and andesite, traced towards the northwest, west and south of the summit. Its lavas are separated from the middle strata by a section that includes lacustrine-swamp deposits and pumice. At the top of the Kydrayvy Volcano, according to geological and morphological features, elements of four craters and an explosion funnel are distinguished, they are differing in age, structure and mode of fumarolic activity.

As a result of the examination, 70 kg samples taken from each geothermal field in the Kudryavy Volcano vent have demonstrated rhenium content up to 387.7 g/t. Topographic survey and analytical study has shown that rare-metal mineralization was formed in high-temperature geothermal fields and covered a total area 7600 m^2, which could exceed the abundance ratio. At the same time, significant rhenium accumulation was revealed in geothermal fields within the temperature ranges of 500 ºC to 613 ºC. Rhenium is found on porous slag, jointed rocks and volcanic breccias.

ICP-MS and NSAM No. 179-X procedures for the element composition measurement have demonstrated high accuracy and expanded uncertainty for assay results. However, ICP-MS method was limited when measuring rhenium in samples with high molybdenum contents. The matrix effect affected the procedure accuracy. Samples should be highly diluted, but it is not possible to avoid molybdenite impact completely to decrease the matrix impact. That’s why we used NSAM No. 179-X. The procedure accuracy was ±0.0018 % wt. to 0.0025 % wt.; expanded uncertainty of procedural results was 0.0068 % wt. to 0.012 % wt. It was achieved by preparing a rhenium (VII) solution, reducing it to rhenium (IV) with a tin (II) chloride solution, converting it into a complex compound, sorbent separation from the solution, diffuse reflectance measurement at 510 nm, and calculation of rhenium (IV) contents using a calibration curve. Silica modified chemically with N-(1,3,4-thiodiazole-2-thiol)-N'-propylurea groups was used as a sorbent.

The designed volcanogenic deposit exploration technology comprised massif crushing, excavation-loading works and transportation.

The geothermal massif of volcanic rocks features different strength even in the rocks of the same type. It is explained by the different rate of fumarole activity (the higher the activity, the higher the material porosity) and temperature zone in the geothermal field. Rock crushing by a mechanical ripper is possible at seismic speed of $V_{ultrasonic}$=1800 m/s and jointing acoustic indicator of R=0.6. This crushing method may be inefficient in Rhenievoye and Kupol fields due to solid inclusions. Taking into account the thermostatic condition for these areas, it is advisable to crush rocks using the electrothermal method at a specific resistance of breakdown path up to 1 kOhm/s, which makes the further mechanical loosening possible.

The studied material demonstrated a very low actual weight of the load in the bucket of the extraction and loading equipment during each digging cycle. The laboratory experiment carried out using the samples revealed a filling rate within the range of 40 % to 60 % of the bucket capacity up to 10 m^3 due to the large rock lumpiness.

Transportation of minerals by a tubular scoop conveyor at an average mineral density of 5.5 t/m^3 may ensure the productivity of 88 m^3/h at a speed of 2.0 m/s and overcome steep slopes and the elevation difference of 900 m between the mining area and the work area.

5. CONCLUSIONS

To sum it up, it is possible to conclude that an annual increase in global rhenium demand requires looking for new deposits. Kudryavy Volcano is a volcanogenic deposit where rhenium mineralization occurs in geothermal fields on a total area of 7200 m^2. Moreover, all rocks of the volcano are divided into three strata of different ages. The lower thickness consists of numerous blocky lava flows of andesite-basalt and andesite composition, exposed at the periphery of the volcano and separated from each other by breccia crusts. The outpouring of lava flows occurred mainly in the western direction; to the east, the spread of lavas was limited by the buildings of the rhyolite dome and Sredny Volcano.

The middle stratum is composed of highly oxidized viscous lavas and an andesitic extrusive dome. The distribution of this stratum is limited by the near-crater part of the volcano.

The upper thickness includes 7 flows of olivine-bearing bipyroxene basaltic andesite and andesite, traced towards the northwest, west and south of the peak. Its lavas are separated from the middle strata by a section that includes lacustrine-swamp deposits and pumice. At the top of the Kudryavy Volcano according to geological and morphological features, elements of four craters and an explosion funnel are distinguished, differing in age, structure and mode of fumarolic activity.

It was found that the maximum rhenium concentration was observed within the temperature range of 500ºC to 613ºC by comparing the results of studying the samples from the volcanic massif and temperature distribution on the geothermal field surface. Mineralization occurred on porous slag, jointed rocks and volcanic breccias.

Rhenium content was measured using a spectrometric procedure that allowed avoiding the impact of high molybdenum contents and attaining high measurement accuracy. The basic concept of such an analysis was to reduce rhenium from oxidation degree 7+ (perrhenate ions) to 4+ in an acidic medium with further sorbent separation from the solution. Silica modified chemically with N-(1,3,4-thiodiazole -2-thiol)-N'-propylurea groups was used as a sorbent.

Exploration of this deposit is complicated by high temperature in geothermal fields up to 900ºC, fumarole gas releases and different strength of rocks of the same type. It was explained by the different rates of fumarole activity (the higher the activity, the higher the material porosity) and temperature zone in the geothermal field. Considering these facts, the technology should comprise the following processes: mechanical loosening, additional loosening using an electrothermal device, excavation by a bucket, and transportation from elevation of 900 m to the working platform via a tubular scoop conveyor.

Acoustic measurement of the rock massif has shown the possibility to mechanically crush the massif at $V_{ultrasonic}$=1 800 m/s and acoustic jointing indicator of R=0.6. However, large temperature differences of geothermal fields Rhenievoye and Kupol formed solid inclusions that required electrothermal crushing with specific resistance of breakdown path of 1 kOhm/cm.

The tubular scoop conveyor productivity could attain 88 m^3/h at a speed of 2.0 m/s at the average density of the mined mineral of 5.5 t/m^3.

Finally, this examination has demonstrated that the geothermal fields of Kudryavy Volcano under study are prospective in terms of rhenium mineralization. The technology for ore mining and transportation to the work site allowed conducting industrial mining operations with acceptable risk.

An increase in the geothermal field temperature could result in a powerful eruption and loss of valuable elements. That is why exploitation of the deposit in Kudryavy Volcano is an urgent task for the near future.

6. REFERENCES

1. Hannington, M.D., "Volcanogenic Massive Sulfide Deposits", ***Treatise on Geochemistry: Second Edition***, Vol. 13, (2013), 463-488. doi: 10.1016/B978-0-08-095975-7.01120-7
2. Hosseinzadeh M., Alizadeh M., "Improvement of the Solvent Extraction of Rhenium from Molybdenite Roasting Dust Leaching Solution using Counter-current Extraction by a Mixer-settler Extractor (TECHNICAL NOTE)", ***International Journal of Engineering, Transactions A: Basics***, Vol. 27, No. 4, (2014), 651-658. (http://www.ije.ir/article_72295.html)
3. Esina N., Minchenko L., Pankrat'ev A., Isaev V., Molchanova Nina, "Electrodeposition of Rhenium onto Oriented Rhenium Foils from Salt Melts", ***Russian Journal of Applied Chemistry***, Vol. 74, (2001), 1461-1464. doi: 10.1023/A:1013784530337
4. Kholmogorov, A., Kononova, O., Kachin, S., Ilyichev, S., Kryuchkov, V., Kalyakina, O., and Pashkov, G., "Ion exchange recovery and concentration of rhenium from salt solutions", ***Hydrometallurgy***, Vol. 51, No. 1, (1999), 19-35
5. Ye Longgang, Ouyang Zhen, Chen Yongming, Liu Shufen, "Recovery of rhenium from tungsten-rhenium wire by alkali fusion in KOH-K_2CO_3 binary molten salt", ***International Journal of Refractory Metals and Hard Materials,*** Vol. 87, (2019), 105-148. doi: 10.1016/j.ijrmhm.2019.105148
6. Donaldson J.G., Hoertel F.W., Cochran A.A., "A preliminary study of vapor deposition of rhenium and rhenium-tungsten", ***Journal of the Less Common Metals***, Vol. 14, (2002), 207-213. doi: 10.1016/0022-5088(68)90207-5
7. Szabo Sandor, Bakos I., "Rhenium deposition onto platinum surface by reduction of perrhenic acid with methanol in hydrochloric acid media", ***Reaction Kinetics and Catalysis Letters***. Vol. 65, (1998). 259-263. doi: 10.1007/BF02475262
8. Iatsenko Gerhardt, N., Palant, A. and Dungan, S., "Extraction of tungsten (VI), molybdenum (VI) and rhenium (VII) by diisododecylamine", ***Hydrometallurgy***, Vol. 55, No. 1, (2000), 1-15
9. Ghorbani, F., Sanati, A., Younesi, H. and Ghoreyshi, A., "The Potential of Date-palm Leaf Ash as Low-cost Adsorbent for the Removal of Pb (II) Ion from Aqueous Solution", ***International Journal of Engineering-Transactions B: Applications***, Vol. 25, No. 4, (2012), 269
10. Alekseev V.I., "Rare Metals in Minerals of Stanniferous Metasomatites from the VerkhneurmiyskyOre Cluster (Far East, Russia)", ***Journal of Mining Institute***, Vol. 232, (2018), 335-340. doi:10.31897/PMI.2018.4.335
11. Solodov N.A., Semenov E.I., "On the Conditions of Finding Rhenium in Nature", ***Geology of Ore Deposits***, Vol. 39, (1997), 106-108.
12. Bradshaw, G.D., Rowins, S.M., Peter, J.M., and Taylor, B.E., "Genesis of the Wolverine deposit, Finlayson Lake district, Yukon: A transitional style of polymetallic massive sulfide mineralization in an ancient con- tinental margin setting", ***The Gangue***, (2003), 1-7
13. Doyle, M.G., and McPhie, J, "Facies architecture of a silicic intrusion- dominated volcanic centre at Highway-Reward, Queensland, Australia", ***Journal of Volcanology and Geothermal Research***, Vol. 99, 2000, 79-96
14. Kailachakov P., Doynikova O., Belousov Peter, Vikentev I., "Unique Rhenium Deposit in the Carboniferous Coal-bearing Sands of the Russian Plate", ***Communication 2. Ore Mineralogy. Lithology and Mineral Resources***, Vol. 55, (2020), 286-315. doi: 10.1134/S0024490220040021
15. Gifkins C., "Textural and Chemical Characteristics of Diagenetic and Hydrothermal Alteration in Glassy Volcanic Rocks", ***Examples from the Mount Read Volcanics, Tasmania. Economic Geology***, Vol. 96, (2001). 973-1002
16. Large, R.R., Gemmell, J.B., Paulick, H., and Huston, D.L., "The alteration box plot: A simple approach to understanding the relationship between alteration mineralogy and lithogeochemistry associated with volcanic-hosted massive sulfide deposits", ***Economic Geology***, Vol. 96, (2001), 957–971
17. Hannington M.D., Bleeker Wouter, Kjarsgaard Ingrid, "Sulfide mineralogy, geochemistry, and ore genesis of the Kidd Creek deposit", ***II The bornite zone. Economic Geology Monograph***, Vol. 10, (1999), 225-266
18. Tkachenko S.I., Taran Y.A., Korzhinsky M.A., Pokrovsky B.G., Steinberg G.S., Shmulovich K.I., "Gas jets of Kudryavy volcano, Iturup Island, Kuril Islands", ***Reports of the Russian Academy of Sciences***, Vol. 325, (1992), 823-828.

19. Znamenskiy V.S., Laputina I.P., Taran Y.A., Yakushev A.I., "Ore deposits from high-temperature gas jets of Kudryavy volcano, Iturup Island, Kuril Islands", ***Reports of the Russian Academy of Sciences***, Vol. 333, (1993), 227-230.

20. Allen, R.L., Weihed, P., and Global VMS Research Project Team, , Global comparisons of volcanic-associated massive sulphide districts, in Blundell, D.J., Neubauer, F., and Von Quadt, A., eds., "The Timing and Location of Major Ore Deposits in an Evolving Orogen", ***Geological Society of London Special Publication***, (2002), 204, 13-37

21. Cheremisina O.V., Cheremisina E.A., Ponomareva M.A., Vilenskaja A.V., "Sorption of coordination compounds of rare-earth elements"**,** ***Journal of Mining Institute***, Vol. 244, (2020), 244-255

22. Allen R.L., Weihed P., "Global VMS Research Project Team, Global comparisons of volcanic-associated massive sulphide districts", The Timing and Location of Major Ore Deposits in an Evolving Orogen: Geological Society of London Special Publication, No. 204, (2000), 13-37

23. United States Department of the Interior, Bureau of Mines, RCRA, "Regulation Impact on Alaska Mineral Development. Prepared by Steffen", Robertson, and Kirsten, Inc., Denver, Colorado, (2010)

24. Evdokimova O.V., Pechishcheva N.V., Shunyaev K.Yu., "Modern methods of rhenium detection", ***Russian Journal of Analytical Chemistry***, Vol. 67, No. 9, (2012), 828-841.

25. Hui Hu, Long Li Sun, Yan Ling Gao, Tian Wang, Yue Fei Zhang, Hui Xiong Wu, Xiao Hui Chen, "High sensitivity thiocyanate spectrophotometric method for determination of perrhenate, an analogue of radioactive pertechnetate, under acidic condition", ***Chemical Papers***, Vol. 73, No. 5, (2019), 1093-1101.

26. Keshavar E., Sadrnezhaad S.K., Mos-hefi Shabestari Z., "Chloroform aided extraction spectrophotometric determination of rhenium using thiocyanate complexing agent", ***Journal of Material Science Technology***, Vol. 21. No. 2, (2005), 239-242.

27. Borisova L.V., Ermakov A.N., "Analytical chemistry of rhenium", Nauka, Moscow, (1974), 161.

28. Znamensky V.S., Korzhinsky M.A., Steinberg G.S., Tkachenko S.I., Yakushev A.I., Laputina I.P., Bryzgalov I.A., Samotoin N.D., Magazina L.O., Kuzmina O.V., Organov Y.I., Rassulov V.A., ChaplyginI.V., "Rheniite, ReS_2 - natural rhenium disulfide from fumaroles of Kudryavy volcano (Iturup Island, Kuril Islands)", ***Proceedings of the Russian Mineralogical Society***, Vol. 5, (2005), 32-34.

29. Klitin A.K., "A brief description of Iturup Island" (2002) https://web.archive.org/web/20050526141324/http://www.sakhalin.ru/Rover/TEXTS/Itur_s.htm

30. Global Volcanism Program, Report on Moyorodake [Medvezhia] (Japan - Under the Control of Russia). V: Sennert S.K. (Ed.), The Weekly Volcanic Activity Report, The Smithsonian Institution & US Geological Survey, (2017)

https://volcano.si.edu/showreport.cfm?doi=GVP.WVAR20170726-290100

31. John Seech. "Medvezhya volcano". Volcano Live http://volcanolive.com/medvezhia.html

32. Layton-Matthews Daniel, Peter Jan, Scott Steven, Leybourne Matthew, "Distribution, Mineralogy, and Geochemistry of Selenium in Felsic Volcanic-Hosted Massive Sulfide Deposits of the Finlayson Lake District, Yukon Territory, Canada", ***Economic Geology***, Vol. 103, (2008), 61-88. doi: 10.2113/gsecongeo.103.1.61

33. E. V. Loginov, O. Iu. Loktiukova, M. E. Melnitskaya, "Calculation of Bench Width for Backhoe Hydraulic Excavators Operated in Sinking Mining Systems", ***Journal of Engineering and Applied Sciences***, Vol. 14, (2019), 6444-6448. doi: 10.36478 / jeasci.2019.6444.6448

34. Scientists: Kudryavy volcano on Iturup is preparing for a powerful eruption (August 25, 2019) https://astv.ru/news/society/2019-08-25-uchenye-vulkan-kudryavyj-na-iturupe-gotovitsya-k-moshnomu-izverzheniyu#comments

35. Gavrishev S.E., Burmistrov K.V., Kornilov S.N., Tomilina N.G., "Evaluation of transportation flow charts with open-pit hoisting systems in open pit/underground mining", ***Gornyi Zhurnal***, Vol. 5, (2016), 41-47

36. Burmistrov K.V., Osintsev N.A., Shakshakpaev A.N., "Selection of Open-Pit Dump Trucks during Quarry Reconstruction", ***Procedia Engineering***, Vol. 206, (2017), 1696-1702

37. Fomin S.I., Ivanov V.V., "Improving the reliability of opencast system for complex structure ore deposits", ***International Journal of Civil Engineering and Technology***, Vol. 9, (2018). 36-43

38. Semenov A.S., "Assessment of Project Risk in the Hierarchical Organization of the Process of Design of Complex Technical Systems", ***World Applied Sciences Journal***, Vol. 30, No. 8, (2014), 1080-1082. doi: 10.5829/idosi.wasj.2014.30.08.14114

39. Fomin S.I., Ivanov V.V., Semenov A.S., Ovsyannikov M.P, "Incremental open-pit mining of steeply dipping ore deposits", ***ARPN Journal of Engineering and Applied Sciences,*** Vol. 15, No. 11, (2020), 1306-1311

40. Yakubovskiy M. M., Isheiskii V. A., "Substantiation the technological schemes for development of hydrolytic lignin's dumps", ***International Journal of Civil Engineering and Technology***, Vol. 9, No.11, (2018), 1640-1646

41. Kovalski E.R., Karpov G.N., Leisle A.V., "Geomechanic models of jointed rock mass", ***International Journal of Civil Engineering and Technology***, No. 6, (2018), 440-448

42. Zuev B.Y., Zubov V.P., Smychnik A.D., " Determination of static and dynamic stresses in physical models of layered and block rock masses", ***Gornyi Zhurnal***, No. 7, (2019), 61-66

43. Zuev B.Y., Zubov V.P., Fedorov A.S., "Application prospects for models of equivalent materials in studies of geomechanical processes in underground mining of solid minerals", ***Eurasian Mining***, No.1, (2019), 8-12

44. Kovalevskyi V.N., Moldovan D.V., Chernobai V.I., "Determination of focal length for cumulative charges with various concavity shapes", ***International Journal of Mechanical Engineering and Technology***, Vol. 8, No. 11, (2017), 1119–1125

45. Sidorov D.V., Ponomarenko T.V., " Estimation methodology for geodynamic behavior of nature-and-technology systems in implementation of mineral mining projects", ***Gornyi Zhurnal***, No. 1, (2020), 49-52. doi: 10.17580/gzh.2020.01.09

46. Mysin A.V., Kovalevskiy V.N., "Creation and Verification of Numerical Model of Explosive Charge Blast in the Ansys Software System, for the Purpose of Substantiating the Optimal Parameters of Drilling and Blasting Operations", E3S Web of Conferences, No. 174, (2020). 1046-1052. doi: 10.1051/e3sconf/202017401046

47. Jakub Gontarz, Jerzy Podgórski, Jonak Józef, Kalita Marek, Siegmund Michał, "Comparison Between Numerical Analysis and Actual Results for a Pull-out Test", (2018)

48. Mashchenko Volodymyr, Khomenko Oleh, Kvasnikov V., "Thermodynamic aspect of rock destruction", Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, Vol. 6, (2020), 25-30. doi: 10.33271/nvngu/2020-1/025

Persian Abstract

چکیده

از رنیوم به طور گسترده ای در صنعت تولید و متالورژی استفاده می شود. امروزه مصرف رنیوم زیاد است ، اما ذخایر بسیار کمی در جهان وجود دارد که در آن استخراج می شود. بنابراین ، هدف از این مطالعه شناسایی مناطق توزیع رنیوم در آتشفشان کودریاوی ، واقع در جزایر خط الراس Kuril (روسیه) است. در این ارتباط ، در طول دوره مزرعه ، ما نمونه هایی از یک توده آتشفشانی به وزن ۷۰ کیلوگرم را برداشتیم ، و همچنین زمینه های زمین گرمایی را با یک فشار سنج بررسی کردیم. تحقیقات آزمایشگاهی شامل مطالعه ترکیب نمونه ها با استفاده از روش طیف سنجی جرمی پلاسما و تجزیه و تحلیل طیف سنجی القایی بود. این مقاله مناطق توزیع کانی سازی رنیوم را تعریف می کند ، نتایج اندازه گیری درجه حرارت زمینه های زمین گرمایی و ترکیب اولیه نمونه های فن آوری را ارائه می دهد. زمین شناسی ، فن آوری توسعه شرح داده شده است ، تجزیه و تحلیل تخریب سنگ ها ، تعیین شده توسط روش صوتی ، مقاومت خاص در برابر شکست در هنگام شل شدن الکترو گرمایی داده شده است. پارامترهای نوار نقاله بیل لوله ای که در آن بهره وری ۸۸ متر مکعب در ساعت حاصل می شود. نتایج به دست آمده امکان شناسایی زمینه های زمین گرمایی امیدوار کننده ، تعیین فناوری توسعه و کمک به مطالعه ذخایر آتشفشانی را فراهم می کند.

IJE TRANSACTIONS A: Basics Vol. 34, No. 4, (April 2021) 1066-1073

International Journal of Engineering

Journal Homepage: www.ije.ir

Investigation of the Influence of Pressures and Proppant Mass on the Well Parameters after Hydraulic Fracturing

V. V. Poplygin*[a], E. E. Pavlovskaia[b]

[a] *Oil and Gas Technologies Department, Mining and Oil Faculty, Perm National Research Polytechnic University, Perm, Russia*
[b] *Centre for Applied Dynamics Research, School of Engineering, University of Aberdeen, Scotland, UK*

PAPER INFO

Paper history:
Received 17 December 2020
Received in revised form 16 January 2021
Accepted 16 February 2021

Keywords:
Hydraulic Fracturing
Productivity Index
Oil Production
Permeability
Pressure

ABSTRACT

This paper considers the implementation of hydraulic fracturing in an oil field located in the Arkhangelsk region in Russia. During the exploitation, the production rates and injectivity of the injection wells in the field were intensively decreased. To increase well flow rates, hydraulic fracturing of the formation was carried out, and the evaluation of the efficiency was performed. Oil production rates after the fracturing increased 3.2 times and the productivity index increased twice. The influence of the geometrical sizes of fractures on the volume of injected proppant was investigated. An increase in the mass of the injected proppant from 2 to 3 tons per 1 meter of formation thickness leads to an average increase in the crack width by 0.5 mm, and the half-length by 40 m. Well work parameters after hydraulic fracturing of a reservoir were obtained as a function of the original parameters of the reservoir. It was observed that there is a sharp decrease in well production rates after fracturing in wells with low bottom hole pressures. When the pressure at the bottom of the well decreases from 60 to 20 MPa, it leads to an average decrease in the crack width by 2 mm, and the half-length of the crack by 50 m. Direct correlation between the well productivity coefficients after fracturing and the values of bottom hole pressures was observed. The optimal conditions for fracturing were identified, which made it possible to significantly increase the efficiency of the operation.

***doi**: 10.5829/ije.2021.34.04a.33*

NOMENCLATURE

P_r	Production rate (t/day)	P_{sat}	Saturation pressure (MPa)
F_r	Flow rate (m^3/day)	PI	Productivity index
P_{bot}	Bottomhole pressures (MPa)		

1. INTRODUCTION

During field development projects for oil fields, especially those characterized by a branched network of cracks, the location of the wells over the area of the deposit needs to be carefully considered as the optimisation of wells location and the interaction of natural and artificial cracks allows increasing the oil production from the reservoir [1].

It is known that the compensation of reservoir volumes of produced fluid by injected volumes of water ensures the consistency of reservoir pressure. As a consequence, relatively stable oil well production rates are maintained, but in the course of time the content of water in the produced fluid increases. However, for waterflooding of low-permeability reservoirs, this condition is often not fulfilled for the following reasons [2]:

- Insufficiently effective transfer of the impact from the injection zone to the selection zone leads to a decrease in reservoir pressure in the drainage area, even if the compensation condition is formally observed.
- The situation is exacerbated by the use of intensive oil production systems based on wells with cracks of considerable length fracturing, horizontal wells with multi-zone fracturing (several cracks from

*Corresponding Author Institutional Email: poplygin@bk.ru (V. V. Poplygin)

different points of the trunk), etc., to develop low-permeability layers.

Successful practical application of hydraulic fracturing technology is not possible without taking into account existing geological and technological limitations [3].

Geological limitations are those which are related to the features of the geological structure of the reservoir and the underlying layers above and below [4, 5, 6, 7]. For example, the proximity of gas or aquifers (less than 10 meters) leads to the need to limit the growth of the crack in height, which automatically leads to the limitation of the maximum volume of the injected proppant, and requires a change in the perforation strategy and a hydraulic fracturing design [8, 9]. Also, geological restrictions can be conditionally referred to the limitations associated with the organization of waterflooding. Decrease in rock brittleness inhibits fracture propagation and lower minimum horizontal stress contrast leads to fracture containment [10, 11].

Technological restrictions include restrictions related to the technical capabilities of the equipment and chemical reagents used [12, 13, 14, 15]. For example, the presence of a restriction on the maximum pressure developed in the process of hydraulic fracturing (primarily due to the capability of wellhead reinforcement and the power of pumping units) can lead to the limitation of the maximum width of the crack.

The highest efficiency of hydraulic fracturing can be achieved by designing its application as an element of the development system, taking into account the well placement system and evaluating their mutual influence in various combinations to treat producing and injection wells in modern software products such as WellFrac, FracPro, Mfrac, StimPt, Petrel, and others. The effect of hydraulic fracturing is not uniformly manifested in the operation of individual wells; therefore, it is necessary to consider not only the increase in the production rate of each well due to hydraulic fracturing, but also the effect of the mutual arrangement of wells, the specific distribution of reservoir heterogeneity, and the energy capabilities of the facility.

In the subject literature, it is noted that such parameters of the technological process of hydraulic fracturing as injection rate of fluid [16], the volume of proppant [17], and applied pressure significantly affect the efficiency of the operation. Nagar et al. [18] noted that the longitudinal cracks in the rocks increase the well production rate more significantly than transversal ones. Therefore, the parameters of fracturing need to be selected, taking into account the geological and physical conditions of the deposit.

The current work aims to investigate different measures to increase the efficiency of hydraulic fracturing in a field focussing on

- analysis of the original conditions and results of the fracturing on the wells;
- definition of the refined criteria of wells selection for carrying out hydraulic fracturing;
- determination of well-working conditions after hydraulic fracturing.

This paper is organised as follows. In the next section, the considered oil field and characteristics of the deposit are described. The field infrastructure and history of operation including the encountered difficulties are detailed in the following section. The performed hydraulic fracturing and analysis of the results are given in section 4 and finally concluding remarks are presented.

2. RESERVOIR INFORMATION

The oil field analysed in this study is located in the southern part of the Nenets Autonomous District of the Arkhangelsk Region. In tectonic terms, the deposit is confined to the central part of the Kolvin megaval. The main oil reserves are confined to the sandstones of the Middle Devonian (D_2st). The oil is light, low-viscosity, the permeability of the collectors is quite high. The oil production of the wells during their operation is significantly reduced, and the water injection is used to enhance the oil recovery. To increase oil production and the volume of injected water, hydraulic fracturing of the formation is actively carried out in the fields. However, during fracturing, well flow rates often decrease significantly and the planned oil production levels are not reached.

The terrigenous deposits of the Old Oskol overhorizon of the Middle Devonian D_2st are represented by sandstones of monomineralic quartz small-medium-grained, medium-coarse-grained and heterogeneous grains, aleurosandstocks, siltstones with various grained, argillites fractured; micro profile is found enriched with fragments of psephitic dimension (gravel-sandstone). Macroscopically the rocks of these sediments are light gray, gray, plot up to a dark gray color, with a brownish tinge, due to uneven oil saturation, strong, dense, porous, stylized, and fractured to varying degrees. The rocks are unevenly carbonized, pyritized, and sulfated**.** In the section of rocks, quartz fine- and medium-fine-grained sandstones are permeable. The basic information about the parameters of the deposit is given in Table 1.

The reservoirs in the deposit have medium porosity, average permeability values, high-gas-saturated and low-viscosity oil. Saturation pressure high values of oil with gas and the initial reservoir pressure are worth being noted.

TABLE 1. D_2st oil reservoir Geological and physical characteristics

Parameters	Value
Depth, m	4129…4084
Type of collector	terrigenous
Effective oil-saturated thickness, m	5.20 -19.5
Coefficient of porosity, fractions of units	0.087- 0.096
Permeability, 10^{-3} μm^2	13.6 - 60.7
Initial reservoir temperature, °C	102
Initial reservoir pressure, MPa	44.3 - 51.8
Viscosity of oil in reservoir conditions, mPa·s	0.69 - 0.81
Oil Density in reservoir, t / m^3	0.690 - 0.705
Oil Density in surface, t / m^3	0.818 - 0.822
Volume factor of oil, fractions of units.	1.36 - 1.42
Oil saturation Pressure, MPa (P_{sat})	18.1 - 23.9
Gas factor, m^3 / t	164.3 - 169.1

Reservoir D_2st (sandstone of the Layer of the Middle Carboniferous Tier) has been put into trial operation in April 2001 by one exploratory well.

In 16 years from the beginning of production, there were 31 wells in the production fund, including 25 units in the current fund. The average daily oil production rate for the wells was 78.1 tons per day. The active injection stock was 3 units.

The period from 6 to 9 years from the beginning of production was characterized by a decrease in the selection of liquid and oil, which was due to a decrease in reservoir pressure in the oil extraction zone, as well as the retirement of four wells into inactivity. The watercut of the production for the same period has remained stable and amounts to 0.4%.

The period from 10 to 11 years from the beginning of production was characterized by a positive dynamics growth in the extraction of liquid and oil, which was due to the introduction and effective operation of the pressure maintenance system.

The operating fund of production wells in 12 years compared to 11 years did not change. But oil production decreased by 30%, the production of liquid by 37%, watercut decreased to 1.9% (the fact of 2011 - 6.7%). The main problems of the development of this facility were:

-Weak coverage by area (the northern and southern parts of the deposit are not involved in the development);

-Anticipatory watering of wells by pumped water through the most permeable interlayers;

-Heterogeneity of the reservoir (the flow rate of the fluid varies significantly in the wells);

-Presence of sites with low reservoir pressure (in the main selection zone). The average weighted reservoir pressure in the extraction zone is 32.3 MPa, which is lower than the initial reservoir pressure by 37.8%. The saturation pressure along the reservoir is 18 MPa;

- availability of a non-working well stock.

In 16 years from the beginning of production, the current oil recovery was 19.2%. The selection from the initial recoverable reserves was 42.75%, with an average watercut for 2016 - 28.3%.

The productive layers of the oilfield were developed with the maintenance of reservoir pressure by flooding. Current and accumulated compensation for the field amounted to 74.5% and 52.5% respectively.

In the period from 10 to 11 years, oil production rate was growing, which is connected with the effective operation of the pressure maintenance system and the results of the measures to increase production rates. But after the 12-year production, the oil production rate began to decline. The increase in watercut for the period under review was 25%.

From 14 to 17 years of production, the following measures were carried out at the field: repair and insulation works were carried out in 4 wells, optimization was performed in 7 wells, 1 reperforation, and 23 fracturings.

Based on wells performance data obtained in the selected field in the results of hydraulic fraction operation, this paper aims to evaluate the efficiency of hydraulic fracturing and develop recommendations for optimizing hydraulic fracturing.

4. HYDRAULIC FRACTURING RESULTS

The main factor that determines the necessity of hydraulic fracturing in wells is the failure to achieve the oil production expected according to the well drilling regulations (known reservoir properties). Low well productivity is usually associated with damage to the bottomhole formation zone caused by the action of the drilling fluid during drilling, blockage of the pores during operation, and the shutdown of wells during underground and overhaul repairs [19, 20]. Low productivity of wells can also be associated with low reservoir properties of the bottomhole formation zone. It is also possible to reduce the permeability in the bottomhole zone due to the deformation of the collectors having high clay content and subject to the highest pressure drops in the bottomhole zone. Therefore, when planning the hydraulic fracturing, an important task is to assess the reason for not obtaining the design parameters of oil production in specific wells.

To increase oil production during years 14 to 17 of operation, hydraulic fracturing of the formation was carried out on 23 wells. As the fracturing fluid, gels with a gellant type DWP-991 and a stitcher type DWP-114 were mainly used. The mass of proppant injected into the reservoir varied from 12 to 78.5 tons. In most cases, proppant type BoroProp was used. The average fracture pressure at fracturing was 52 MPa. The average values of the parameters of the formed fracture cracks were as follows: crack opening - 1.8 mm, height - 15 m, half-length - 164 m.

The main indicators of the wells before the activities and in the first month after the fracturing are given in Table 2.

In the field, oil production rates after the fracturing increased 3.2 times, productivity index increased 2 times. The average watercut of wells increased from 29.88% to 30.81%.

The wells were operated at average reservoir pressures mainly slightly below the gas saturation pressure. The bottomhole pressure after fracturing was 0.6 to 1.5 relative to saturation pressure. The specific gas content of the oil in the free gas in the reservoir at the walls of the wells at the lowest bottomhole pressures reached 0.5. At such gas content values, the relative permeability of rocks over oil can decrease several-fold.

TABLE 2. The main parameters of wells operation (average values)

Parameters	Before frac	After frac
Flow rate, m^3 / day	32.6	88.35
Oil production rate, t / day	15.88	52.22
Watercut,%	29.88	30.81
Reservoir pressure, MPa	27.42	27.42
Bottom-hole pressure, MPa	17.45	23.68
Productivity index, m^3/(MPa·day)	3.01	6.29

To increase the flow rate in the well 609 located in the area of the injection well (Figure 1) hydroulic fracturing was conducted. However, the reservoir pressure in its area was reduced by 20% of the initial value. During hydraulic fracturing, it was very important that the fracture does not propagate towards the injection well. Therefore, to ensure the success of the process, the hydraulic fracturing was planned using modern software products with the construction of the conductivity profile. With the decrease in Biot coefficient, stress inversion region will decrease significantly [21]. The pump flow increases and the crack deflection distance increases [22]. The pre-fracturing parameters were as follows: Reservoir pressure– 39.6 MPa; Bottom pressure– 14.6 MPa; Productivity index - 0.96 m^3/(day·MPa); Flow rate - 23.9 m^3/day; Oil production rate - 18.8 t/day; Watercut - 0%. Parameters after frac: Bottom pressure - 14.8 MPa; Productivity index - 24.8 m^3/(day·MPa); Flow rate - 69 m^3/day; Oil production rate - 40.7 t/day; Watercut - 10%.

The results showed that appropriate choice of fracturing technology and parameters has been made demonstrating improve production rate.

5. ANALYSIS OF RESULTS AND DISCUSSION

Hydraulic fracturing allows to increase the well production rate by creating highly permeable channels. The size of the channels depends on the mass of the proppant injected and the pressure in the reservoir.

The effectiveness of hydraulic fracturing can be traced by the change in the productivity index. However, the productivity index immediately after hydraulic fracturing can increase and then sharply decrease due to deformations of productive formations, gas release and deposits of wax and salts.

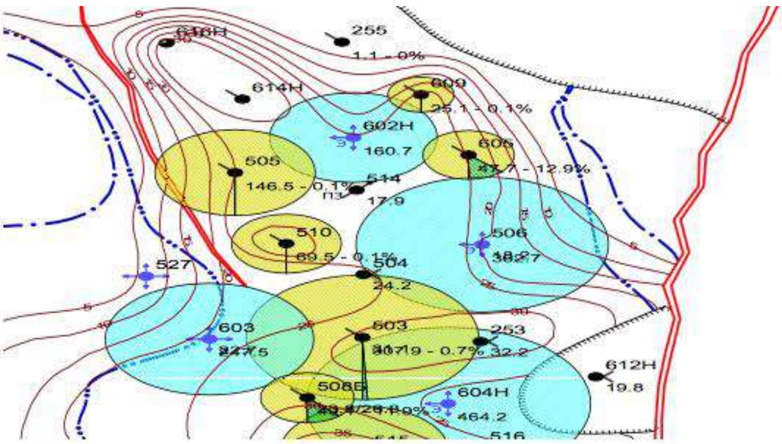


Figure 1. Well 609 location plan

The dependence of the productivity index on the ratio of bottomhole pressure and saturation pressure was analyzed and the following equation of the productivity index after fracturing depending on the relative bottomhole pressure was obtained:

$$PI = 4.4\ P_{bot}/P_{sat} + 0.14 \quad (1)$$

For the considered interval P_{bot}/P_{sat} 0.6-1.5, the correlation coefficient of the linear dependence was 0.95.

The duration of the effect from the fracturing was 16 months. After fracturing, productivity index and production rate of the well began to decline rapidly. For 16 months after hydraulic fracturing, the flow rate of the well has decreased from 1.2 (relative to the original) to 0.5. This can be associated with low bottomhole pressure in wells, which values are in the range of 0.6 from the saturation pressure.

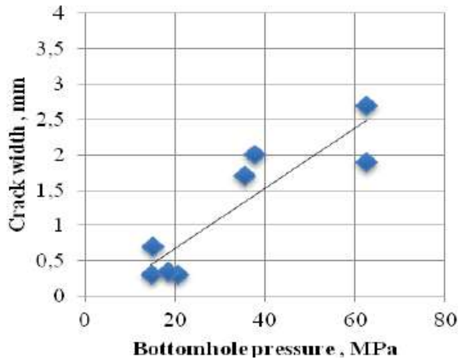
Dependence of crack width and half-length of cracks of the field D_2st object after the fracturing on the bottomhole pressure is shown in Figure 2. It can be noted that there is a clear relationship between the bottomhole pressure after fracturing and the well productivity factor. This ligature is caused by deformations of rocks [23] with a decrease in reservoir pressure and release of free gas at bottom pressure below the saturation pressure.

Using the value of reservoir pressure, the production wells can be divided into two groups: with the formation pressure in the drainage area greater than the saturation pressure and with the formation pressure less than the saturation pressure. Wells with reservoir pressure above the saturation pressure are characterized by the largest increase in oil flow rates and the duration of the effect from the event. It is noted that the dynamics of productivity coefficients in the operation of wells after fracturing depends significantly on bottomhole pressure. With a decrease in bottomhole pressure and an increase in depression in the reservoir, the coefficients of well productivity decreased, well flow rates neither increased or slightly decreased. With the increase in bottomhole pressure after fracturing, the productivity factor and oil rates increased.

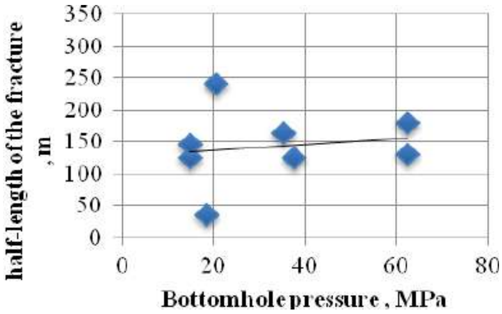
Pressure values also affect the size of cracks occurring during the fracturing process (Figures 2 and 3). The number of cracks depends on how effectively hydraulic fracturing was performed. In this case, the width of the cracks increases with increasing reservoir pressure and saturation pressure ratios [24].

With a decrease in saturation pressure and an increase in reservoir pressure, the ratio on the whole increases, the size of the crack increases, and as a result the effective hydraulic fracturing is performed [25, 26].

The dependence of the geometric dimensions of the formed cracks on the mass of the injected proppant per



(a)



(b)

Figure 2. Dependence of crack width (a) and half-length of cracks (b) after fracturing on bottomhole pressure

1 meter of perforated thickness is shown in Figure 3. As the concentration of the wedging element increases, the crack width on the average also increases. It can be said that as the proppant mass is increased, the size of the cracks increases, which leads to a good result. In some cases, the formation of plugs is possible, so the determination of the proppant concentration should be made taking into account all other factors, such as the injection rate, the volume of the fracturing fluid, the quality of the wedging element [27, 28]. As the volume of injected proppant increases, the openness and half-length of the cracks increases. The cost of the fracturing work is determined by the feasibility and environmental impact assessment. Fracture indicators are usually evaluated depending on the length of cracks. The cost of work on hydraulic fracturing increases with the length of the crack. The income curve without the cost of hydraulic fracturing has a maximum at some length. The optimal value of the proppant mass per meter of perforated thickness for the deposit is 2 t / m. This volume of the proppant ensures the most optimal cracks both from the point of view of technological efficiency and costs [29, 30].

The longer the crack is, the larger the inflow from the formation to the fracture is. But at the same time, the

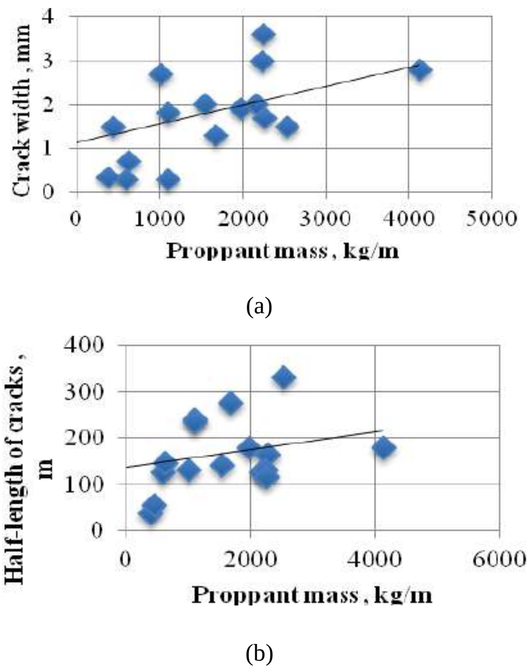


Figure 3. Dependence of crack width (a) and half-length of cracks (b) after fracturing on proppant mass per 1 meter of perforated thickness.

crack becomes narrower, and the cross-sectional area decreases, and for some values of its conductivity, the crack begins to play the role of a chimney [31]. As a result, the pressure drop in the crack increases. Accordingly, the effective pressure drop between the fracture and the formation decreases. That is, with the growth in the fracture length there are two opposing factors: an increase in the area of inflow to the fracture, increasing the production rate of the well, and a decrease in the effective pressure drop between the fracture and the formation, which reduces the production rate of the well. Consequently, for a fixed fracture volume, there is a ratio of the half-length of the fracture and its width at which the flow rate of the well is maximized.

The intensive development system implemented in the field is far from realizing its potential due to the lack of efficiency of pressure maintenance system. In such conditions, the medium- and long-term dynamics of the production rates, especially for wells with hydraulic fracturing, is determined not by their productivity, but by the possibilities of compensating the volume of fluids taken out of the formation. Based on the results of the analysis it is possible to give a presumed efficiency estimate for other wells in which the action is only planned.

The conducted study confirms that the selection of candidate wells for the application of the hydraulic fracturing technology has to be carried out according to the following criteria [32, 33]:

- current reservoir pressure is not lower than the gas saturation pressure;
- effective oil-saturated thickness is not less than 3.5 m;
- the thickness of the dense bridge is more than 3 m;
- the tightness of the production column and the absence of streaked flows;
- a satisfactory state of the cement stone in the perforation interval.

According to the results obtained the optimal conditions for hydraulic fracturing in the field under consideration are: maintaining bottomhole and reservoir pressure above the saturation pressure, specific proppant flow rate - about 2 t/m, proppant type – BoroProp.

6. CONCLUSION

The paper presents the results of hydraulic fracturing study conducted in the field located in the southern part of the Nenets Autonomous District of the Arkhangelsk Region in Russia. When analyzing the work of the well stock, it was noted that well flow rates were intensively decreasing due to the deformation of rocks and release into the free phase of dissolved gas in the oil. Large depth of wells with relatively low permeability values led to the fact that wells after construction and repair were hard to put into operation.

In order to improve the oil production, hydraulic fracturing has been actively used in the field. When analyzing the technological efficiency of hydraulic fracturing, a significant relationship between the geometric parameters of cracks and the energy state of the formation in the borehole region was revealed. An increase in the mass of the injected proppant from 2 to 3 tons per 1 meter of formation thickness leads to an average increase in the crack width by 0.5 mm, and the half-length by 40 m. Analogous relationships also exist in the evaluation of the productivity and production rates of wells after fracturing. When the pressure at the bottom of the well decreases from 60 to 20 MPa, it leads to an average decrease in the crack width by 2 mm, and the half-length of the crack by 50 m. The optimal conditions for hydraulic fracturing in the considered field were achieved by maintaining bottomhole and reservoir pressure above the saturation pressure, using a specific proppant flow rate at about 2 t/m for the selected proppant type (BoroProp). Based on the obtained dependencies, it is possible to predict the results of hydraulic fracturing in the field. For example, if we carry out a fracturing operation taking into account the above recommendations, we will increase the oil production rate by about 20%.

7. ACKNOWLEDGEMENT

The researche was supported by a grant from the Russian Science Foundation (Project No.19-79-10034).

8. REFERENCES

1. Zhou D., He P., "Major Factors affecting simultaneous frac results", ***SPE Production and Operations Symposium, Proceedings***, Vol. 2015, (2015), 532-545. doi.org/SPE-173633-MS
2. Kalantarias A., Schulze K., Storz J., Burmester C., Kuenckeler S., You Z., Badalyan A., Bedrikovetsky P., "Produced water re-injection and disposal in low permeable reservoirs", ***Journal of Energy Resources Technology, Transactions of the ASME***. Vol. 141, (2019), Article number 072905. https://doi.org/10.1115/1.4042230
3. Hooker J. N., Larson T. E., Eakin A., Laubach S. E., Eichhubl P., Fall A., "Marrett Randall Fracturing and fluid flow in a sub-décollement sandstone; or, a leak in the basement", ***Journal of the Geological Society***, Vol. 172, (2015), 428-442. https://doi.org/10.1144/jgs2014-128
4. Zhao H., Zhen Li, Zhu C., Ru Z., "Reliability analysis models for hydraulic fracturing", ***Journal of Petroleum Science and Engineering***, Vol. 162, (2018), 150-157. https://doi.org/10.1016/j.petrol.2017.12.048
5. Kachanov M., Prioul R., Jocker J., "Incremental linear-elastic response of rocks containing multiple rough fractures: Similarities and differences with traction-free cracks", ***Geophysics***, Vol. 75, (2010), https://doi.org/10.1190/1.3268034.
6. Zhuang L., Kim K.Y., Jung S.G., Diaz M., Min K.-B., "Effect of Water Infiltration, Injection Rate and Anisotropy on Hydraulic Fracturing Behavior-of Granite", ***Rock Mechanics and Rock Engineering***, Vol. 2, (2018), 575-589 https://doi.org/10.1007/s00603-018-1431-3.
7. Hu Q.-J., Li M.X., Qiao M.P., Liu C.C., Liu S.Q., Fan B., Sang G.J., Yan L., "Analysis of key geologic factors of fracturing effect of CBM wells for high-rank coal in Southern Qinshui Basin//Meitan Xuebao", ***Journal of the China Coal Society***, Vol. 42, (2017), 1506-1516.
8. Fan M., McClure J., Han Y., Li Z., Chen C., "Interaction between proppant packing, reservoir depletion, and fluid flow in hydraulic fractures", ***Proceedings of the Annual Offshore Technology Conference***, Vol. 3, (2017), 1913-1931. https://doi.org/10.4043/27907-MS
9. Li M., Guo P.-J., Liang L., Li X., "Hydraulic fracturing characteristics of heterogeneous rock with hard inclusion distributed", ***Yantu Lixue/Rock and Soil Mechanics***, Vol. 37, (2016), 3130-3136.
10. Wan X., Rasouli V., Damjanac B., Pu H., "Lattice simulation of hydraulic fracture containment in the North Perth Basin", ***Journal of Petroleum Science and Engineering***, Vol. 188, (2020), Article number 106904. https://doi.org/10.1016/j.petrol.2020.106904
11. Nagaso M., Mikada H., Takekawa J., "The role of rock strength heterogeneities in complex hydraulic fracture formation – Numerical simulation approach for the comparison to the effects of brittleness", ***Journal of Petroleum Science and Engineering,*** Vol. 172, (2019), 572-587. https://doi.org/10.1016/j.petrol.2018.09.046
12. Zhang L., Zhou J., Braun A., Han Z., "Sensitivity analysis on the interaction between hydraulic and natural fractures based on an explicitly coupled hydro-geomechanical model in PFC2D", ***Journal of Petroleum Science and Engineering***, Vol. 167, (2018), 638-653. https://doi.org/10.1016/j.petrol.2018.04.046
13. Zhao Y.-L., Li N.-Y., Zhang L.-H., Zhang R.-H., "Productivity analysis of a fractured horizontal well in a shale gas reservoir based on discrete fracture network model", ***Journal of Hydrodynamics***, Vol. 31, (2019), 552-561.
14. Teng B., Li H.A., "A semi-analytical model for characterizing transient flow behavior of reoriented refractures", ***Journal of Petroleum Science and Engineering***, Vol. 177, (2019), 921-940. https://doi.org/10.1016/j.petrol.2019.02.005
15. Sun J., Wu S., Deng J., "An analytical model for fracture initiation in elasto-plastic soft formations based on stress path analysis", ***Journal of Petroleum Science and Engineering***, Vol. 176, (2019), 906-925. https://doi.org/10.1016/j.petrol.2019.01.064
16. Gil I., Sebastian H., Wendte J., Lee P., Patel H., Cadwallader S., Sun T., Merletti G., Fry M., "The combination of innovative completion hardware and 3D non-planar fracture/reservoir simulation in shale completion optimization", 50th US Rock Mechanics/ Geomechanics Symposium 2016, Vol. 4, (2016), 3100-3107.
17. Shelley R.P.E., Shah K., Guliyev N., Mohammadnejad A., Sheludko S., "Is pumping large volume sand frac treatments sustainable?" Proceedings- SPE Annual Technical Conference and Exhibition, Vol. 2015, (2015), 1907-1920. https://doi.org/10.2118/174863-MS
18. Nagar A., Tibbles R., Guha R., Chidambaram P., Sinha A., Pathak S., Agarwal P., Misra A., Ranjan A., Ranjan V., Tiwari S., Dwivedi N., "The advantages of longitudinally fracturing the Barmer Hill porcellinite - A case study of successfully fractured horizontal wells in Barmer Basin", Rajasthan, India. SPE Annual Technical Conference and Exhibition. ATCE 2016; Dubai; United Arab Emirates; (2016). https://doi.org/10.2118/181376-MS
19. Dvoynikov M. V., Nutskova M. V., Blinov P. A., "Developments Made in the Field of Drilling Fluids by Saint Petersburg Mining University", ***International Journal of Engineering, Transactions A: Basics,*** Vol. 33, (2020), 702-711. DOI: 10.5829/IJE.2020.33.04A.22
20. Ezzatneshan E., "Oil Reservoirs Classification Using Fuzzy Clustering", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 30, (2017), 1391-1400. DOI.: 10.5829/idosi.ije.2017.30.09c.12
21. Hu Y.,Wang Q., Zhao J., Chen S., Zhang G., Wang S., Wang Z., Fu C., "Study on induced stress of hydraulic fracturing in fractured-porous elastic media", ***Energy Science and Engineering***, Vol. 8, (2020). 3314-3332. https://doi.org/10.1002/ese3.721
22. Lin J., Xu C., Yang J. "Influence of pump flow rate on the deflection of hydraulic fracturing cracks in longitudinal flume cutting" ***Meitan Xuebao/Journal of the China Coal Society***, Vol. 45, (2020), 2804-2812.
23. Pratap N. S., Tobias M. M., "Diffusion in deformable porous media: Incompressible flow limit and implications for permeability estimation from microseismicity", ***Geophysics,*** Vol. 85, (2020). https://doi.org/10.1190/geo2019-0510.1
24. Alex H.**,** Mark Z., "The relation between stimulated shear fractures and production in the Barnett Shale: Implications for unconventional oil and gas reservoirs", ***Geophysics***, Vol. 84, (2019), 461-469. https://doi.org/10.1190/geo2018-0545.1.
25. Lyu W., Zeng L., Zhou S., Du X., Xia D., Liu G., Li J., Weng J., "Natural fractures in tight-oil sandstones: A case study of the Upper Triassic Yanchang Formation in the southwestern Ordos Basin", ***AAPG Bulletin***, Vol. 103, (2019), 2343-2367.https://doi.org/10.1306/0130191608617115
26. Zeng Q., Liu W., Yao J., "Hydro-mechanical modeling of hydraulic fracture propagation based on embedded discrete fracture model and extended finite element method", ***Journal of Petroleum Science and Engineering***, Vol. 167, (2018), No. 8, 64-77. https://doi.org/10.1016/j.petrol.2018.03.086
27. Xie P., Gao Z., Li C., Fan T., Grace M.K., Wei D., Wang S., Machel H.G., "Conductivity of hydraulic fracturing in tight carbonate intra-platform shoal reservoirs", ***Journal of Petroleum Science and Engineering***, Vol. 189, (2020), Article number 106976. https://doi.org/10.1016/j.petrol.2020.106976

28. Makedonska N., Karra S., Viswanathan H.S., Guthrie G.D. ,“Role of interaction between hydraulic and natural fractures on production”, ***Journal of Natural Gas Science and Engineering***, Vol. 82, (2020), Article number 103451. https://doi.org/10.1016/j.jngse.2020.103451

29. Mordvinov V.A., Poplygin V.V., Chalov S.V., “Change in the productivity of the obtaining bore holes with the development of the deposits of oil with the high gas saturation”, ***Neftyanoe khozyaystvo - Oil Industry***, Vol. 8, (2010), 104-106.

30. Mordvinov V.A., Poplygin V.V., Erofeev A.A., “Influence of gas reservoir and deformation on the performance of wells after fracturing”, ***Neftyanoe khozyaystvo- Oil Industry***, Vol. 10, (2012), 102-103.

31. Tan P., Pang H., Zhang R., Jin Y., Zhou Y., Kao J., Fan M., “Experimental investigation into hydraulic fracture geometry and proppant migration characteristics for southeastern Sichuan deep shale reservoirs”, ***Journal of Petroleum Science and Engineering***, Vol. 184, (2020), Article number 106517. https://doi.org/10.1016/j.petrol.2019.106517

32. Raspopov A.V., Kondratiev S.A., Sharafeev R.R., Novokreschennykh D.V., Drozdov S.A., “Experience of hydraulic fracturing in oil fields of the perm region, the Komi republic and the Nenets autonomous district”, ***Neftyanoe Khozyaystvo- Oil Industry*****.** Vol. 8, (2019), 48-51. https://doi.org/10.24887/0028-2448-2019-8-48-51

33. Voevodkin V.L., Chertenkov M.V. “New Technologies in LUKOIL: From simple to complicated”, ***Neftyanoe Khozyaystvo- Oil Industry*****,** Vol. 8, (2019), 62-66. https://doi.org/10.24887/0028-2448-2019-8-62-66

Persian Abstract

در این مقاله اجرای شکستگی هیدرولیکی در یک میدان نفتی واقع در منطقه ارخانگلسک در روسیه مورد بررسی قرار می گیرد. در حین بهره برداری ، میزان تولید و تزریق چاه های تزریق در مزرعه به شدت کاهش می یابد. برای افزایش سرعت جریان چاه ، شکست هیدرولیکی سازند انجام شد و ارزیابی کارایی انجام شد. نرخ تولید روغن پس از شکستگی 3.2 برابر و شاخص بهره وری دو برابر افزایش یافت. تأثیر اندازه های هندسی شکستگی بر روی حجم غلاف تزریق شده بررسی شده است. افزایش جرم غلاف تزریق شده از 2 به 3 تن در هر 1 متر ضخامت تشکیل منجر به افزایش متوسط عرض ترک 0.5 میلی متر و نیمه طول 40 متر می شود. پارامترهای خوب کار پس از شکست هیدرولیک مخزن به عنوان تابعی از پارامترهای اصلی مخزن بدست آمد. مشاهده شد که پس از شکستگی در چاههای با فشار سوراخ پایین ، کاهش شدیدی در میزان تولید چاه وجود دارد. هنگامی که فشار در پایین چاه از 60 به 20 مگاپاسکال کاهش می یابد ، منجر به کاهش متوسط عرض ترک 2 میلی متر و نیمه طول ترک 50 متر می شود. همبستگی مستقیم بین ضرایب بهره وری چاه پس از شکستگی و مقادیر فشار سوراخ پایین مشاهده شد. شرایط بهینه برای شکستگی مشخص شد ، که امکان افزایش قابل توجه کارایی عملیات را فراهم کرد.

AIMS AND SCOPE

The objective of the International Journal of Engineering is to provide a forum for communication of information among the world's scientific and technological community and Iranian scientists and engineers. This journal intends to be of interest and utility to researchers and practitioners in the academic, industrial and governmental sectors. All original research contributions of significant value focused on basics, applications and aspects areas of engineering discipline are welcome.

This journal is published in three quarterly transactions: Transactions A (Basics) deal with the engineering fundamentals, Transactions B (Applications) are concerned with the application of the engineering knowledge in the daily life of the human being and Transactions C (Aspects) - starting from January 2012 - emphasize on the main engineering aspects whose elaboration can yield knowledge and expertise that can equally serve all branches of engineering discipline.

This journal will publish authoritative papers on theoretical and experimental researches and advanced applications embodying the results of extensive field, plant, laboratory or theoretical investigation or new interpretations of existing problems. It may also feature - when appropriate - research notes, technical notes, state-of-the-art survey type papers, short communications, letters to the editor, meeting schedules and conference announcements. The language of publication is English. Each paper should contain an abstract both in English and in Persian. However, for the authors who are not familiar with Persian, the publisher will prepare the latter. The abstracts should not exceed 250 words.

All manuscripts will be peer-reviewed by qualified reviewers. The material should be presented clearly and concisely:

- *Full papers* must be based on completed original works of significant novelty. The papers are not strictly limited in length. However, lengthy contributions may be delayed due to limited space. It is advised to keep papers limited to 7500 words.
- *Research* notes are considered as short items that include theoretical or experimental results of immediate current interest.
- *Technical notes* are also considered as short items of enough technical acceptability with more rapid publication appeal. The length of a research or technical note is recommended not to exceed 2500 words or 4 journal pages (including figures and tables).

Review papers are only considered from highly qualified well-known authors generally assigned by the editorial board or editor in chief. Short communications and letters to the editor should contain a text of about 1000 words and whatever figures and tables that may be required to support the text. They include discussion of full papers and short items and should contribute to the original article by providing confirmation or additional interpretation. Discussion of papers will be referred to author(s) for reply and will concurrently be published with reply of author(s).

INSTRUCTIONS FOR AUTHORS

Submission of a manuscript represents that it has neither been published nor submitted for publication elsewhere and is result of research carried out by author(s). Presentation in a conference and appearance in a symposium proceeding is not considered prior publication.

Authors are required to include a list describing all the symbols and abbreviations in the paper. Use of the international system of measurement units is mandatory.

- On-line submission of manuscripts results in faster publication process and is recommended. Instructions are given in the IJE web sites: www.ije.ir-www.ijeir.info
- Hardcopy submissions must include MS Word and jpg files.
- Manuscripts should be typewritten on one side of A4 paper, double-spaced, with adequate margins.
- References should be numbered in brackets and appear in sequence through the text. List of references should be given at the end of the paper.
- Figure captions are to be indicated under the illustrations. They should sufficiently explain the figures.
- Illustrations should appear in their appropriate places in the text.
- Tables and diagrams should be submitted in a form suitable for reproduction.
- Photographs should be of high quality saved as jpg files.
- Tables, Illustrations, Figures and Diagrams will be normally printed in single column width (8cm). Exceptionally large ones may be printed across two columns (17cm).

PAGE CHARGES AND REPRINTS

The papers are strictly limited in length, maximum 6 journal pages (including figures and tables). For the additional to 6 journal pages, there will be page charges. It is advised to keep papers limited to 3500 words.

Page Charges for Papers More Than 6 Pages (Including Abstract)	
For International Author ***	**$55 / per page**
For Local Author	**100,000 Toman / per page**

AUTHOR CHECKLIST

- Author(s), bio-data including affiliation(s) and mail and e-mail addresses).
- Manuscript including abstracts, key words, illustrations, tables, figures with figure captions and list of references.
- MS Word files of the paper.

International **J**ournal of **E**ngineering, P.O. Box 13145-1659, Tehran, I.R. Iran
Materials and Energy Research Center, 5 Ahuramazda St., Alvand Ave., Argentina Sq., Tehran, Islamic Republic of Iran
www.merc.ac.ir - www.emerc.ac.ir