A comparative analysis of metaheuristic algorithms in fuzzy modelling for phishing attack detection

Noor Syahirah Nordin[1], Mohd Arfian Ismail[2], Tole Sutikno[3], Shahreen Kasim[4], Rohayanti Hassan[5], Zalmiyah Zakaria[6], Mohd Saberi Mohamad[7]
[1,2]Faculty of Computing, College of Computing and Applied Sciences, Universiti Malaysia Pahang, Pahang, Malaysia
[3]Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[4]Faculty of Computing Science and Information Technology, Universiti Tun Hussein Onn, Johor, Malaysia
[5,6]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
[7]Department of Genetics and Genomics, College of Medical and Health Sciences, United Arab Emirates University, United Arab Emirates

Article Info

Article history:

Received May 11, 2021
Revised Jul 9, 2021
Accepted Jul 13, 2021

Keywords:

Cyber security
Fuzzy logic
Metaheuristic algorithm
Phishing attack detection
Phishing websites dataset
Website phishing dataset

ABSTRACT

Phishing attack is a well-known cyber security attack that happens to many people around the world. The increasing and never-ending case of phishing attack has led to more automated approaches in detecting phishing attack. One of the methods is applying fuzzy system. Fuzzy system is a rule-based system that utilize fuzzy sets and fuzzy logic concept to solve problems. However, it is hard to achieve optimal solution when applied to complex problem where the process of identify the fuzzy parameter becomes more complicated. To cater this issue, an optimization method is needed to identify the parameter of fuzzy automatically. The optimization method derives from the metaheuristic algorithm. Therefore, the aim of this study is to make a comparative analysis between the metaheuristic algorithms in fuzzy modelling. The study was conducted to analyse which algorithm performed better when applied in two datasets: website phishing dataset (WPD) and phishing websites dataset (PWD). Then the results were obtained to show the performance of every metaheuristic algorithm in terms of convergence speed, and four metrics including accuracy, recall, precision, and f-measure.

Corresponding Author:

Mohd Arfian Ismail
Faculty of Computing
College of Computing and Applied Sciences
Universiti Malaysia Pahang
26600 Pekan, Pahang, Malaysia
Email: arfian@ump.edu.my

1. INTRODUCTION

Phishing attack is one of the simple and dangerous cyber security threats. It is an action of stealing other person private's information for the attacker's own benefits. The person who did this crime is called a phisher. The aim focuses on identity theft, financial benefits, defame, damaging an organization's reputation and sometimes to gain popularity among the phishers itself. According to phishing activity trends report in third quarter of 2020 by the anti-phishing working group (APWG), software-as-a-service (SaaS) and webmail sites are the most targeted industry for phishing with 31.4 percent followed by financial institution with 19.2 percent [1]. For payment, social media, e-commerce and retail industry, all three industries hold record of 13.4%, 12.6% and 7.2% respectively. Another 16.2 percent comes from another industry including cloud storage with 2.1%, telecom and logistics 3.2% and 4.2% respectively. Meanwhile, other than the mentioned industry holds a percentage of 6.7% in the most targeted industry of the phishing attack. Figure 1 illustrates the division of the most-targeted industries of phishing in third quarter of 2020.

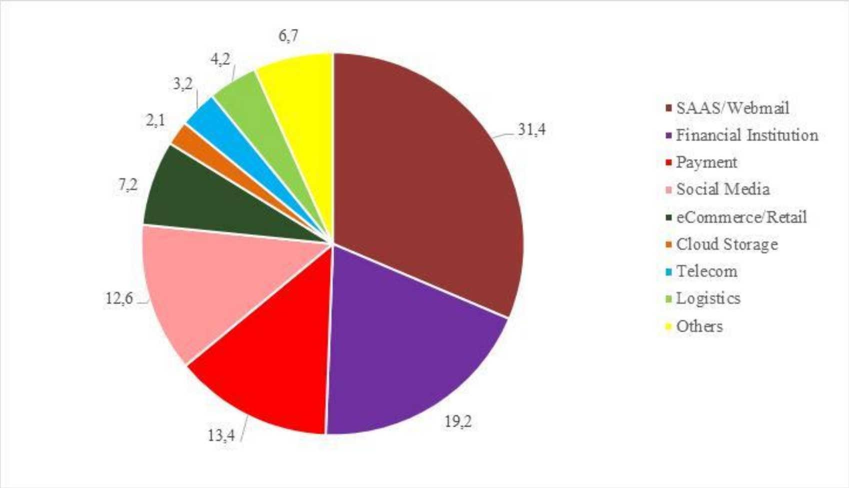

Figure 1. Targeted industry for phishing attack in 3rd quarter of 2020 [1]

Phishing attack can be viewed as one of the classification problems due to the fact that the data i.e. websites need to be categorized into phishing or legitimate websites. Therefore, there exist various methods or approaches proposed by many researchers in order to detect the phishing attack. It is found that the classification method is a promising method to be applied in this case such as fuzzy systems [2]-[4], artificial neural network (ANN) [5], [6], support vector machine (SVM) [7]-[9], and decision tree algorithm (DTA) [10], [11]. ANN able to work with incomplete data but the method is hard to predict the model [12]. SVM uses kernel in the model to learn the functions however the results are commonly difficult to interpret and understand by the decision maker [13]. Meanwhile, DTA is a very simple method to interpret and easy to implement. The drawbacks of this method are it is not being bale to cope with the lost data and the tree must be rebuilt every time a new sample is added to find the solutions. [14], [15]. In consequence, fuzzy system is a good method that can deal with inaccurate and incomplete issues [3], [16].

In fuzzy system, one of the processes is to identify the fuzzy parameter named fuzzy rules and membership functions. This process is called fuzzy modelling. The construction of fuzzy system becomes complicated when it is applied to a complex issue hence the results produced by the system are not guarantee optimal in terms of the system accurateness. Therefore, an optimization method is needed to automate the process of identifying the fuzzy parameter in the fuzzy system. Based on the observation in the previous works done by other researchers, applying metaheuristic algorithm is a well-liked approach that has been used since ages for many purposes [17], [18]. As instances, genetic algorithm (GA) [19], [20]. differential evolution algorithm (DE) [21], [22], particle swarm optimization (PSO) [23]-[25], butterfly optimization algorithm (BOA) [26], [27], teaching-learning-based optimization (TLBO) [28], [29], harmony search algorithm (HSA) [30], [31], and gravitational search algorithm (GSA) [32]-[34]. For that reason, a comparative analysis of metaheuristic algorithms based on the performances is carried out in this study. Seven algorithms were proposed to determine the best algorithm in fine-tuning the parameter in the fuzzy system.

Next section is the detail explanation of each category will be viewed in the next section followed by the research method section. In that section, data collection, experimental design and performance measurement are stated thoroughly. Results and discussion will be in the next part before this paper is wrapped with a conclusion of overall study.

2. METAHEURISTIC ALGORITHM

Metaheuristic algorithm can be categorized into four categories which are evolution-based method, swarm-based method, human-based method, and physics-based method [35], [36] as shown in the Figure 2. Every example of metaheuristic algorithm that fall in each category will be compared and analyzed in the experiment phase. The reason behind the chosen metaheuristic algorithm for each category simple because they are widely used and has shown effective result in phishing and fuzzy modelling [37]-[39].

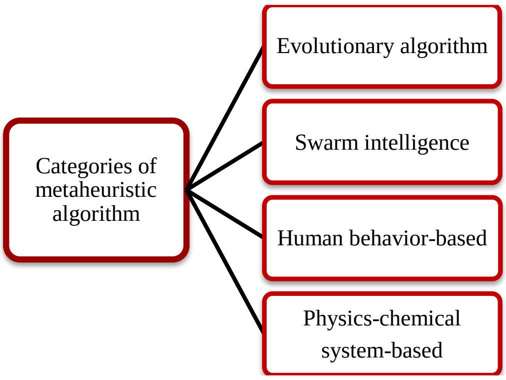

Figure 2. Categories of metaheuristic algorithm

2.1. Evolution-based method

2.1.1. Genetic algorithm

GA was developed that follow the principle of the biological evolution process and was discovered in 1975. It is a robust search approach to solve a wide range problem developed by Holland [40]. The process involves are reproduction, crossover, and mutation. The best genes are called parent chromosomes while the new chromosomes obtained are known as child chromosomes. The procedure of GA is shown in Figure 3.

```
Initialize population
while stopping criteria not met do
    Evaluate fitness
    Selection
    Crossover
    Mutation
end while
Output the best individual found
```

Figure 3. Pseudo code of GA

In [37], the authors propose a method to combine fuzzy and GA, where GA used as optimization method in the fuzzy system. By using dissolution and sintering process in the manufacture of aluminum foams, fuzzy-GA can describe the inherent uncertainties. As a result, the proposed method is a promising tool to be used in manufacturing process. Another work that implemented fuzzy and GA is from [20]. In their work, the routing in dynamic environments were optimize by using fuzzy and GA. Fuzzy logic reduce the time consume to reach destination, meanwhile GA was utilized to tune fuzzy rules table to reduce the travelled distance.

2.1.2. Differential evolution

Storn and Price [11] proposed the DE algorithm which based on population that similar with GA. The concept of this algorithm is quite similar with GA where it is inspired by the species' evolution lived in this world. There are three operators in this algorithm: mutation, crossover, and selection. In DE process, the new vectors (new generation of population) is generated by mutation and crossover process, then the selection process take place to determine whether the new generated vectors would survives in the next generation or not. Figure 4 presents the pseudo code of DE.

```
Initialize population
while stopping criteria not met do
        Mutation
        Crossover
        Selection
end while
Output the best individual found
```

Figure 4. Pseudo code of DE

A paper in [41] has presented a new adaptive DE based on fuzzy inference system where fuzzy was used to tune the mutation factor in DE. The proposed method proved to have a better result than the other method mentioned in the study. Other than that, [38] has propose a method name MODE-FM, which was a multi-objective DE and combine with fuzzy dynamic mutation factor. The mutation factor was tuned by fuzzy. The tuning process was done using count generation and diversity of population. This intended to overcome the lack of DE. As a result, the proposed method seems to have a promising result compared with previous work. In product line design (PLD), fuzzy and GA also has been applied in this field. Fuzzy logic used to calculate the parameter automatically lead to DE that settings-free and shows a promising result [21].

2.2. Swarm-based method

2.2.1. Particle swarm optimization

Kennedy and Eberhart were the first person who developed the PSO algorithm in their work [14]. It is a population-based stochastic optimization method that mimics the social behavior of birds flocking and fish schooling. There are certain parameters in the original version of PSO called control parameters. The parameters involved are acceleration coefficients, velocity clamping-limit, swarm size and maximum number of iterations. Many modifications have been made to improve the performance of the standard PSO. The pseudo code of PSO can be reviewed in Figure 5.

```
Objective function f(x), x = (x_1, x_2, ..., x_n)
Initialize the parameters
while stopping criteria not met do
        for (i=1) in population do
                Calculate fitness function, f
                Update personal and global best of each particle
                Update velocity, v
                Update position, p
        end for i
Output the best particle
end while
```

Figure 5. Pseudo code of PSO

PSO was used as optimization method to tune the membership function in the fuzzy system as it is hard to determine the parameter manually. Therefore, Nurmaini and Setianingsih [42] has proposed a method using PSO and fuzzy to control the position of differential drive mobile robot (DDMR) and resulting to faster time for the robot to reach steady-state condition. In phishing area, [39] used PSO to weight various features in website in order to reach higher accuracy in the result produced. The method was proposed to enhance the phishing website detection process where PSO able to differentiate between the features. By using dataset from UCI machine learning repository, the proposed method shows an outstanding performance compared to the previous methods.

2.2.2. Butterfly optimization algorithm

The recent nature inspired algorithm called BOA was introduced by Arora in his work in [43]. In order to perform optimization, butterflies act as the search agent in BOA. There are three phases in the algorithm: (i) initialization phase, (ii) iteration phase and (iii) final phase. In each iteration, all butterflies will be evaluated by calculating its fitness function before generating the fragrance using (1).

$$f = cI^{a} \tag{1}$$

where f is the fitness function where it supposed to attract other butterflies with their fragrance. Meanwhile c is the sensory modality, I is the variation of butterfly and a denote as power exponent parameter depends on the sensory modality. Then, the iteration will continue until the termination criteria satisfied. Figure 6 presents the steps of BOA in pseudo code.

A works by Fan *et al.* [44] has introduced a new improved BOA to enhance the searching process and the iteration capability in solving numerical optimization problem. The authors have used self-adaption method in BOA named SABOA that applied new iteration, updating strategy and new fragrance coefficient in the basic BOA. As a result, the proposed method gives advantages in terms of precision value, iterative speed and simple structure compared with other algorithm mentioned in the paper. Other than that, BOA was also used as optimization method to tune the fuzzy parameter automatically in fuzzy system [45]. In evaluating the proposed method, the phishing website dataset that obtained from repository of UCI machine learning was used. The result of the proposed method shows a promising and competitive result compared to other metaheuristic algorithm mentioned in the paper.

```
Objective function f(x), x = (x_1, x_2, ..., x_dim), dim dim = no. of dimensions
Generate initial population of n butterflies x_i = (i = 1, 2, ..., n)
Stimulus Intensity I, at x_i is determined by f(x_i)
Define sensor modality, power exponent and switch probability
while stopping criteria not met do
        for each butterfly in population do
                Calculate fragrance of the butterfly using (1)
        end for
        Find the best butterfly
        for each butterfly in population do
                Generate a random number of rand from [0, 1]
                if rand < p then
                        Move towards best butterfly solution
                else
                        Move randomly
                end if
        end for
        Update the value of a
end while
Output the best solution found
```

Figure 6. Pseudo code of BOA

2.3. Human-based method

2.3.1. Teaching-learning-based optimization

TLBO is one of the modern heuristic optimization algorithms that simulates a scenario of teaching and learning between teacher and student in a classroom environment. It is proposed by Rao *et al.* in 2011 and demonstrates a good performance in solving various problems [19]. This algorithm has two fundamental parts: teacher phase and learner phase. Students are considered as population and student with the best fitness considered as a teacher based on the grades obtained in the evaluation process. In the teacher phase, students seek knowledge from the teacher where they have a role to upgrade the students' knowledge level. The TLBO flow is simplified in Figure 7.

```
Initialize learner and evaluate
X_teacher will represent best teacher
Calculate the mean, X_mean of learners
while stopping criteria not met do
        for all learners do
                Calculate the teaching factor, TF
                Update all learners
        end for
        Evaluate new learners
        Accept new learners if better
        for all learners do
                Another learner selects randomly
                Update the learners
        end for
        Accept new learner if better
        Update teacher and mean
end while
```

Figure 7. Pseudo code of TLBO

An adaptive method has been proposed in [29] name ATLBO in solving the process to generate mixed strength t-way test suite problem. The researcher improves TLBO by applying adaptive selection and fuzzy to keep the searching process in balance. In another work, TLBO was also combined with mutated fuzzy adaptive PSO to classify the breast cancer disease. PSO parameters were tuned by the fuzzy system while the hybridization of TLBO and PSO able to solve the optimization problem [46].

2.3.2. Harmony search

HS is a population algorithm developed based on the process of finding the perfect state of harmony in music used by musician [47]. Three main components in HS are usage of harmony memory, pitch adjusting and randomization. The first component is very important to ensure that the best harmonies will be chosen as the new harmony memory. Then, the pitch adjustment will be determined by a pitch band-width b_{range} and r_{pa} which represents pitch adjusting rate. The process of adjusting the pitch can be simplified in (2).

$$x_{new} = x_{old} + b_{range} * \varepsilon \tag{2}$$

where x_{new} will be the new pitch after pitch adjustment step and x_{old} is the existing pitch in the harmony memory while ε represents random number generator in range of [-1, 1]. The pseudo code of HS is shown in Figure 8.

```
Objective function f(x), x = (x_1, x_2, ..., x_dim)^T
Define initial harmonics (real number arrays)
Define Pitch Adjusting Rate (r_pa), pitch limits and bandwidth
Define harmony memory accepting rate (r_accept)
while (t < Max number of iterations)
        Generate new harmonics by accepting best harmonics
        Adjust pitch to get new harmonics (solutions)
        if (rand > r_accept), then
                Choose an existing harmonic randomly
        else if (rand > r_pa) then
                Adjust the pitch randomly within limits
        else
                Generate new harmonics via randomization
        end if
        Accept the new harmonics (solutions) if better
end while
Find the current best solutions
```

Figure 8. Pseudo code of HS

In a machining system, a fuzzy model and HS was developed to increase the ability and performance of the system. The system was categorized as failure, repair and vacation in fuzzy numbers and HS used to handle the cost optimization problems in the machine repair model. The proposed method has been proved affective to be applied in the machining system [31]. In another field which is power system security, fuzzy-HS was also applied to achieve in the problem of optimal power flow to find the best solution. The study proposed that the adjustment of HS parameter was handled by fuzzy logic system. As a result, the proposed method able to improve the security problems in power system [48].

2.4. Physics-based method

2.4.1. Gravitational search algorithm

Rashedi, NezamabadiPour and Saryazdi have proposed a new optimization algorithm in 2009 named gravitational search algorithm (GSA) based on Newton’s law of gravitation and motion of individuals in nature [26]. The searching agent in GSA is the object with a specific mass. The interaction between agents is considered as the global movement in the algorithm. The calculation of the active, passive and inertia mass are as (3).

$$M_i(t), P_i(t), I_i(t) \propto f_i(t) \tag{3}$$

where $M_i(t), P_i(t), I_i(t)$ represents active, passive and inertia mass respectively while $f_i(t)$ is the objective value of *i* at the time *t*. The gravitational constant G is a function of time where G_0 that will be presented in (4) acts as the initial value.

$$G(t) = G(G_0, t) \tag{4}$$

The pseudo code of GSA can be viewed in Figure 9.

```
Generate initial population randomly
while stopping criteria not met do
        Evaluate fitness
        Update active (M), passive (P) and inertia (I) using (3)
        Update gravitational constant, G(t) using (4)
        Calculate forces
        Update accelerations, a
        Update velocities, v
        Update positions, p
end while
```

Figure 9. Pseudo code of GSA

Zhao *et al.* [49] has proposed a new method based in enhance GSA and fuzzy c-means algorithm for oilfield system to detect the unreliable data in the power system. The works done by applying enhanced GSA to search the measurement data and fuzzy c-means to classify data before using concentration of similarity (CoS) metrics clustering validity to determine the unreliable data. The result produced was more accurate in terms of the solution quality. Besides that, GSA and fuzzy were also used for energy storage system (ESS) [50]. The function of GSA was to maximize the fuzzy satisfaction function and to determine the scheduling in active distribution system (ADS). For the phishing website detection, GSA was used as feature selection tool that can eliminate the unnecessary feature. By using dataset from PhishTank and Yandex Search API, the method that proposed by them outperformed other methods in feature subsets selection [51].

3. METHOD

Datasets from the University of California, Irvine (UCI) machine learning repository was used in the experiment. The datasets from this database are a high-quality and trusted data. It can be accessed from http://archive.ics.uci.edu/ml/. The datasets are well understood and can be used freely by everyone for research purposes. Thereby, the datasets used are related to phishing websites: website phishing dataset (WPD) and phishing websites dataset (PWD). WPD can be accessed from https://archive.ics.uci.edu/ml/datasets/Website+Phishing# while PWD from https://archive.ics.uci.edu/ml/datasets/Phishing+Websites#.

In order to test the metaheuristic algorithms, experiments were executed by using k-fold cross validation techniques for predicting the classification algorithm performance. This method is one of the popular methods as it is simple and easy to understand. Moreover, seven algorithms were compared in terms of accuracy, precision, recall and f-measure. These four measurements are the most well-known metrics used in the evaluation process and it is also suitable to be used in this study to make a comparative analysis between the methods mentioned. It is possible to formulate all these measurements as (5), (6), (7) and (8).

$$Accuracy = \frac{(TP+TN)}{no\ of\ data} \tag{5}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{6}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{7}$$

$$F - measure = \frac{(2 \times precision \times recall}{(precision + recall)} \tag{8}$$

where *TP* is true positive, *TN* is true negative, *FN* is false negative, *FP* false positive and number of data is the number of data that has been tested. *TP* is when the cases are predicted yes and the result is yes and *TN* is when the cases are predicted no and the result is no. Meanwhile, *FN* is when the cases are predicted no and the result is yes and it is otherwise for *FP*.

4. RESULTS AND DISCUSSION

The experimental results were collected and presented in this section. The collected results were showed different reading in term of accuracy, recall, precision, and f-measure. In addition, all results were compared by using convergence graph where it measures the convergence rate of each method. Moreover, radar chart was used in comparing accuracy, recall, precision, and f-measure. The results of statistical test also recorded to show the significance difference between each method. Figure 10 and 11 plotted the graph of the fitness per generation to see the convergence speed of every method. The plotted results were recorded from one single run for every method. In both graphs, it shown that BOA converged faster than the other methods where it started with the value of 90 at generation of 1.

Based on the result of the convergence graphs in Figure 10 and Figure 11, BOA has outperformed other metaheuristic methods. BOA has fast convergence rate because of the employed random walk and elitism in the algorithm. The parameter of switch probability decides whether to move to the best butterfly who emits more fragrance or to perform a random walk in the population thus contribute to faster convergence rate of BOA [43]. Other than that, for the radar chart in WPD, the measurement result shows different outcome. The accuracy value of every method shows not much difference where all of them obtained high accuracy value. The next measurement is recall and BOA has obtained highest value compared to other methods. The result of precision value also shows not much difference for every method where all methods obtained the value in the range of 0.78 to 0.89. The last measurement is f-measure where BOA and TLBO obtained the highest value followed by other five algorithms. The summarization of the result of all methods can be seen at Table 1 and Figure 12 to view the radar chart.

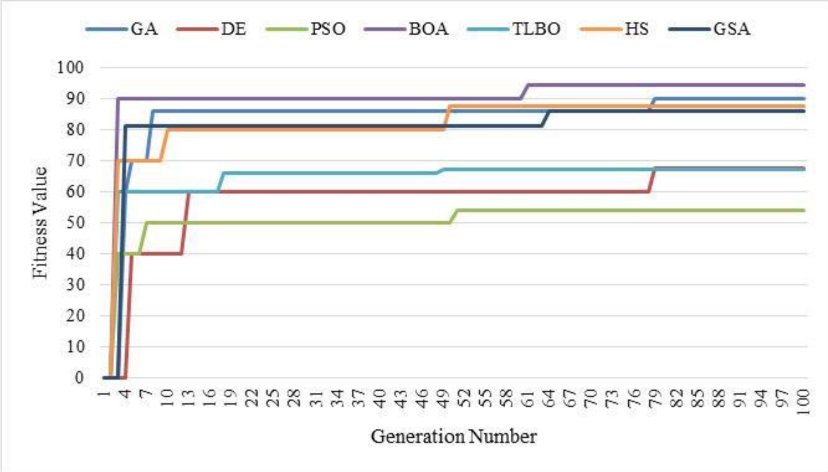

Figure 10. Convergence graph of WPD dataset

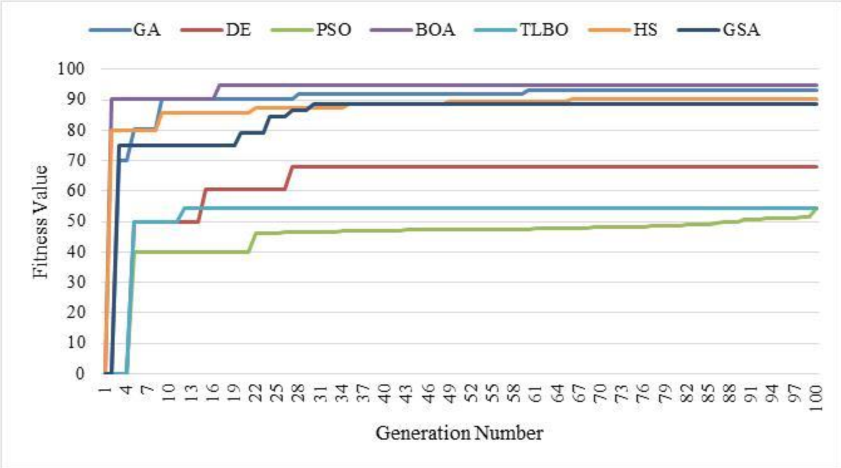

Figure 11. Convergence graph of PWD dataset

Table 1. Comparison results of all methods for WPD

Method	Accuracy	Recall	Precision	F-measure
GA	0.925	0.355	0.878	0.701
DE	0.94	0.512	0.83	0.899
PSO	0.91	0.415	0.788	0.825
BOA	**0.962**	**0.891**	**0.899**	**0.98**
TLBO	0.918	0.5	0.8	**0.98**
HS	0.887	0.82	0.791	0.978
GSA	0.948	0.299	0.873	0.596

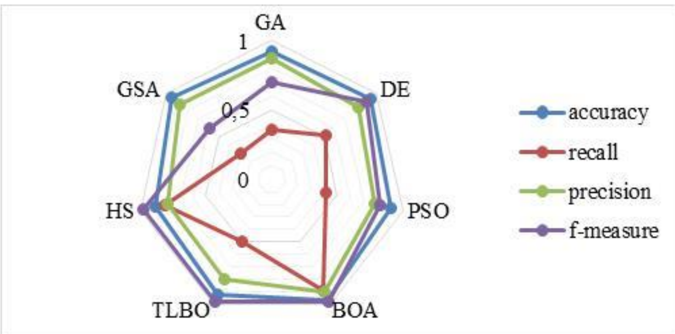

Figure 12. Radar chart of the performance for all methods for WPD

As for the PWD dataset, it can be seen that the best result is obtained by BOA where it indicates highest result for every measurement calculated. For the accuracy and recall value, BOA's result is the highest followed by other algorithms. As for the precision value, BOA and GSA obtained highest value followed by HS where it shows only small different between these three methods. Other than that, the result of f-measure shows that almost all methods have produced high value except for GA and GSA. BOA shows high value in every aspect because of the fragrance attenuation that allows the algorithm to search the solution space efficiently [52]. Table 2 lists the value of each measurement from the seven methods tested while Figure 13 illustrates the performance of the seven methods mentioned earlier in a radar chart. Moreover, 10 independent runs were performed in the statistical test determining whether the results produced by all methods differ statistically from each other. The paired t-test and Wilcoxon signed-rank test were used. Table 3 indicates the fitness value in both statistical tests for WPD and PWD respectively.

Table 2. Comparison results of all methods for PWD

Method	Accuracy	Recall	Precision	F-measure
GA	0.857	0.393	0.756	0.783
DE	0.667	0.498	0.512	0.908
PSO	0.714	0.455	0.756	0.904
BOA	**0.923**	**0.802**	**0.874**	**0.993**
TLBO	0.75	0.746	0.756	0.904
HS	0.801	0.702	0.845	0.972
GSA	0.78	0.299	0.873	0.593

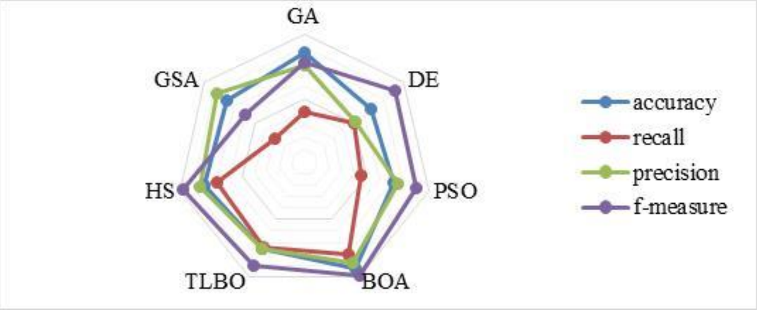

Figure 13. Radar chart of the performance for all methods for PWD

Table 3. Statistical test result of fitness value

Method	WPD		PWD	
	p-value t-test	p-value Wilcoxon signed-rank test	p-value t-test	p-value Wilcoxon signed-rank test
GA vs DE	0.00001	0.00008	0.00001	0.00008
GA vs PSO	0.00001	0.00008	0.00001	0.00008
GA vs BOA	0.00001	0.00008	0.00001	0.00008
GA vs TLBO	0.00001	0.00008	0.00001	0.00008
GA vs HS	0.000027	0.00132	0.00001	0.00008
GA vs GSA	0.00001	0.00034	0.00001	0.00012
DE vs PSO	0.00001	0.00512	0.00001	0.001
DE vs BOA	0.00001	0.00008	0.00001	0.00008
DE vs TLBO	0.000013	0.00116	0.00001	0.0006
DE vs HS	0.00001	0.00008	0.00001	0.00008
DE vs GSA	0.00001	0.00008	0.00001	0.00008
PSO vs BOA	0.00001	0.00008	0.00001	0.00008
PSO vs TLBO	0.00001	0.0001	0.000117	0.00168
PSO vs HS	0.00001	0.00008	0.00001	0.00008
PSO vs GSA	0.00001	0.00008	0.00001	0.00008
BOA vs TLBO	0.0001	0.00008	0.00001	0.00008
BOA vs HS	0.00001	0.00008	0.00001	0.00008
BOA vs GSA	0.00001	0.00008	0.00001	0.00008
TLBO vs HS	0.00001	0.00008	0.00001	0.00008
TLBO vs GSA	0.00001	0.00008	0.00001	0.00008
HS vs GSA	0.001525	0.04338	0.006506	0.00252

From the observation, all p-values for paired t-test and Wilcoxon signed-rank test were smaller than the value of α in both datasets. Therefore, it can be concluded that all methods have significance different with each other.

5. CONCLUSION

Metaheuristic algorithm is one of the optimization methods that can be utilized in the fuzzy modelling for phishing attack detection. Seven different methods were discussed and comparative analyses have been made in this study which are genetic algorithm (GA), differential evolution (DE), particle swarm optimization (PSO), butterfly optimization algorithm (BOA), teaching-learning-based optimization (TLBO), harmony search (HS), and gravitational search algorithm (GSA). The methods were compared through analysing the convergence rate, accuracy, precision, recall, and, f-measure value in the radar chart and the p-value in the statistical test. From the result, it can be seen that BOA outperformed other six metaheuristic algorithms in both datasets.

ACKNOWLEDGEMENTS

This study was supported by Postgraduate Research Grants Scheme (PGRS) with vot No. PGRS200396. Some of this work also was supported/funded by the Ministry of Higher Education under Fundamental Research Grant Scheme (FRGS/1/2018/ICT01/UTM/02/4).

REFERENCES

[1] APWG, "Phishing Activity Trends Report 3rd Quarter 2020," 2020.
[2] P. Barraclough and G. Sexton, "Phishing Website Detection Fuzzy System Modelling," *2015 Science and Information Conference (SAI)*, 2015, pp. 1384-1386, doi: 10.1109/SAI.2015.7237323.
[3] H. Chapla, R. Kotak, and M. Joiser, "A Machine Learning Approach for URL Based Web Phishing Using Fuzzy Logic as Classifier," *Proc. Fourth Int. Conf. Commun. Electron. Syst. (ICCES 2019)*, no. Icces, 2019, pp. 383-388, doi: 10.1109/ICCES45898.2019.9002145.
[4] K. N. M. Kumar and K. Alekhya, "Detecting Phishing Websites using Fuzzy Logic," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 5, no. 10, pp. 2413-2417, 2016.
[5] E. Zhu, C. Ye, D. Liu, F. Liu, F. Wang, and X. Li, "An Effective Neural Network Phishing Detection Model Based on Optimal Feature Selection," *2018 IEEE Intl Conf Parallel Distrib. Process. with Appl. Ubiquitous Comput. Commun. Big Data Cloud Comput. Soc. Comput. Networking, Sustain. Comput. Commun.*, pp. 781-787, 2018.
[6] E. Zhu, Y. Ju, Z. Chen, F. Liu, and X. Fang, "DTOF-ANN : An Artificial Neural Network phishing detection model based on Decision Tree and Optimal Features," *Applied Soft Computing*, vol. 95, Oct. 2020, doi: 10.1016/j.asoc.2020.106505.
[7] R. Karnik, "Support Vector Machine Based Malware and Phishing Website Detection," *Int. J. Comput. Technol.*, vol. 3, no. 5, pp. 295-300, 2016.
[8] A. A. Orunsolu, A. S. Sodiya, and A. T. Akinwale, "A predictive model for phishing detection," *J. King Saud Univ. - Comput. Inf. Sci.*, 2019, doi: 10.1016/j.jksuci.2019.12.005.
[9] W. Niu, X. Zhang, G. Yang, Z. Ma, and Z. Zhuo, "Phishing emails detection using CS-SVM," in *Proceedings - 15th IEEE International Symposium on Parallel and Distributed Processing with Applications and 16th IEEE International Conference on Ubiquitous Computing and Communications, ISPA/IUCC 2017*, 2018, pp. 1054-1059, doi: 10.1109/ISPA/IUCC.2017.00160.
[10] R. Bawm, "Phishing Attack Detection using Machine Learning Classification Techniques," *2020 3rd Int. Conf. on Intelligent Sustainable Systems (ICISS)*, 2020, pp. 1125-1130, doi: 10.1109/ICISS49785.2020.9315895.
[11] X. Yang, L. Yan, B. Yang, and Y. Li, "Phishing Website Detection Using C4. 5 Decision Tree," *Int. Conf. Inf. Technol. Manag. Eng.*, vol. 2, no. Itme, 2017, pp. 119-124, doi: 10.12783/dtcse/itme2017/7975.
[12] R. P. Ferreira, *et al.*, "Artificial Neural Network for Websites Classification with Phishing Characteristics," *Sci. Reseacrh Publ.*, vol. 7, pp. 97-109, 2018, doi: 10.4236/sn.2018.72008.
[13] S. Karamizadeh, S. M. Abdullah, M. Halimi, J. Shayan, and M. J. Rajabi, "Advantage and drawback of support vector machine functionality," *I4CT 2014 - 1st Int. Conf. Comput. Commun. Control Technol. Proc.*, 2014, pp. 63-65, doi: 10.1109/I4CT.2014.6914146.
[14] A. Soofi and A. Awan, "Classification Techniques in Machine Learning: Applications and Issues," *J. Basic Appl. Sci.*, vol. 13, no. September, pp. 459-465, 2017, doi: 10.6000/1927-5129.2017.13.76.
[15] A. Yasin and A. Abuhasan, "An Intelligent Classification Model for Phishing Email Detection," *Int. J. Netw. Secur. Its Appl.*, vol. 8, no. 4, pp. 55-72, 2016, doi: 10.5121/ijnsa.2016.8405.
[16] A. A. Zuraiq and M. Alkasassbeh, "Review: Phishing Detection Approaches," *2019 2nd International Conference on new Trends in Computing Sciences (ICTCS)*, 2019, pp. 1-6, doi: 10.1109/ICTCS.2019.8923069.
[17] M. Abdel-basset, L. Abdel-fatah, and A. K. Sangaiah, "Metaheuristic Algorithms: A Comprehensive Review," *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, pp. 185-231 2018, doi: 10.1016/B978-0-12-813314-9.00010-4.

[18] T. Dokeroglu, E. Sevinc, T. Kucukyilmaz, and A. Cosar, “A survey on new generation metaheuristic algorithms,” *Comput. Ind. Eng.*, vol. 137, no. May, p. 106040, 2019, doi: 10.1016/j.cie.2019.106040.
[19] A. P. Plerou, E. Vlamou, and V. Papadopoulos, “Fuzzy Genetic Algorithms : Fuzzy Logic Controllers and Genetics Algorithms,” *Glob. J. Res. Anal.*, vol. 5, no. 11, pp. 1-5, 2017.
[20] S. A. F. Aghda and M. Mirfakhraei, “Improved routing in dynamic environments with moving obstacles using a hybrid Fuzzy-Genetic algorithm,” *Future Generation Computer Systems*, vol. 112, pp. 250-257, 2020, doi: 10.1016/j.future.2020.05.024.
[21] S. Tsafarakis, K. Zervoudakis, and A. Andronikidis, “Fuzzy self-tuning differential evolution for optimal product line design,” *European J. of Oprl. Res.*, vol. 287, no. 3, pp. 1161-1169, 2020, doi: 10.1016/j.ejor.2020.05.018.
[22] S. Brindha and S. Miruna Joe Amali, “A robust and adaptive fuzzy logic based differential evolution algorithm using population diversity tuning for multi-objective optimization,” *Eng. Appl. Artif. Intell.*, vol. 102, no. March, p. 104240, 2021, doi: 10.1016/j.engappai.2021.104240.
[23] A. Afzal and M. K. Ramis, “Multi-objective optimization of thermal performance in battery system using genetic and particle swarm algorithm combined with fuzzy logics,” *J. Energy Storage*, vol. 32, no. September, p. 101815, 2020, doi: 10.1016/j.est.2020.101815.
[24] T. Salameh, P. P. Kumar, E. T. Sayed, M. A. Abdelkareem, H. Rezk, and A. G. Olabi, “Fuzzy modeling and particle swarm optimization of Al2O3/SiO2 nanofluid,” *International Journal of Thermofluids*, vol. 10, 2021, doi: 10.1016/j.ijft.2021.100084.
[25] L. Zou, “Design of reactive power optimization control for electromechanical system based on fuzzy particle swarm optimization algorithm,” *Microprocess. Microsyst.*, vol. 82, no. January, 2021, doi: 10.1016/j.micpro.2021.103865.
[26] L. Sen Tan, Z. Zainuddin, and P. Ong, “Wavelet neural networks based solutions for elliptic partial differential equations with improved butterfly optimization algorithm training,” *Appl. Soft Comput. J.*, vol. 95, p. 106518, 2020, doi: 10.1016/j.asoc.2020.106518.
[27] P. Maheshwari, A. K. Sharma, and K. Verma, “Energy efficient cluster based routing protocol for WSN using butterfly optimization algorithm and ant colony optimization,” *Ad Hoc Networks*, vol. 110, no. 2020, p. 102317, 2021, doi: 10.1016/j.adhoc.2020.102317.
[28] B. K. Sahu, S. Pati, P. K. Mohanty, and S. Panda, “Teaching-learning based optimization algorithm based fuzzy-PID controller for automatic generation control of multi-area power system,” *Appl. Soft Comput. J.*, vol. 27, pp. 240-249, 2015, doi: 10.1016/j.asoc.2014.11.027.
[29] K. Z. Zamli, F. Din, S. Baharom, and B. S. Ahmed, “Fuzzy adaptive teaching learning-based optimization strategy for the problem of generating mixed strength t -way test suites,” *Eng. Appl. Artif. Intell.*, vol. 59, no. December 2016, pp. 35-50, 2017, doi: 10.1109/ICSEngT.2017.8123413.
[30] S. M. Mousavi, S. Abdullah, S. T. A. Niaki, and S. Banihashemi, “An intelligent hybrid classification algorithm integrating fuzzy rule-based extraction and harmony search optimization: Medical diagnosis applications,” *Knowledge-Based Syst.*, vol. 220, p. 106943, 2021, doi: 10.1016/j.knosys.2021.106943.
[31] R. Kumar, M. Jain, S. Singh, and A. Assad, “Fuzzy modeling and harmony search optimization for machining system with general repair, standby support and vacation,” *Appl. Math. Comput.*, vol. 361, pp. 858-873, 2019, doi: 10.1016/j.amc.2019.05.053.
[32] D. Kar, M. Ghosh, R. Guha, R. Sarkar, and L. Garcia-hernandez, “Fuzzy mutation embedded hybrids of gravitational search and Particle Swarm Optimization methods for engineering design problems,” *Eng. Appl. Artif. Intell.*, vol. 95, no. July, p. 103847, 2020, doi: 10.1016/j.engappai.2020.103847.
[33] B. González, F. Valdez, P. Melin, and G. Prado-Arechiga, “Fuzzy logic in the gravitational search algorithm for the optimization of modular neural networks in pattern recognition,” *Expert Syst. Appl.*, vol. 42, no. 14, pp. 5839-5847, 2015, doi: 10.1016/j.eswa.2015.03.034.
[34] D. Pelusi, R. Mascella, L. Tallini, J. Nayak, B. Naik, and A. Abraham, “Neural network and fuzzy system for the tuning of Gravitational Search Algorithm parameters,” *Expert Systems with Applications*, vol. 102, pp. 234-244, 2018, doi: 10.1016/j.eswa.2018.02.026.
[35] Z. Wei, C. Huang, X. Wang, T. Han, and Y. Li, “Nuclear Reaction Optimization: A Novel and Powerful Physics-Based Algorithm for Global Optimization,” *IEEE Access*, vol. 7, pp. 1-9, 2019, doi: 10.1109/ACCESS.2019.2918406.
[36] R. Rajakumar, P. Dhavachelvan, and T. Vengattaraman, “A Survey on Nature Inspired Meta-Heuristic Algorithms with its Domain Specifications,” *2016 International Conference on Communication and Electronics Systems (ICCES)*, 2016, pp. 1-6, doi: 10.1109/CESYS.2016.7889811.
[37] G. Salvatore, S. Guarino, O. Giannini, F. Tagliaferri, S. Venettacci, and F. Trovalusci, “Aluminium foam production control by by using a combined algorithm model,” *Procedia CIRP*, vol. 88, pp. 503-508, 2019, doi: 10.1016/j.procir.2020.05.087.
[38] A. Jamali, R. Mallipeddi, M. Salehpour, and A. Bagheri, “Multi-objective differential evolution algorithm with fuzzy inference-based adaptive mutation factor for Pareto optimum design of suspension system,” *Swarm Evol. Comput.*, vol. 54, no. November 2019, p. 100666, 2020, doi: 10.1016/j.swevo.2020.100666.
[39] W. Ali, “Particle Swarm Optimization-Based Feature Weighting for Improving Intelligent Phishing Website Detection,” vol. 8, 2020, doi: 10.1109/ACCESS.2020.3003569.
[40] J. H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, Cambridge, United States: MIT Press, 1975, doi: 10.7551/mitpress/1090.001.0001.
[41] M. Salehpour, A. Jamali, and A. Bagheri, “A new adaptive differential evolution optimization algorithm based on fuzzy inference system,” *Engineering Science and Technology, an International Journal*, vol. 20, no. 2, pp. 587-597, 2017, doi: 10.1016/j.jestch.2017.01.004.

[42] S. Nurmaini and F. Setianingsih, "Enhancement of The Fuzzy Control Response with Particle Swarm Optimization in Mobile Robot System," *2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 2018, pp. 73-78, doi: 10.1109/ICECOS.2018.8605221.

[43] S. Arora and S. Singh, "Butterfly optimization algorithm : a novel approach for global optimization," *Soft Comput.*, vol. 23, no. 3, pp. 715-734, 2019, doi: 10.1007/s00500-018-3102-4.

[44] Y. Fan, J. Shao, G. Sun, and X. Shao, "A Self-Adaption Butterfly Optimization Algorithm for Numerical Optimization Problems," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2993148.

[45] N. S. Nordin, M. Ismail, and N. Omar, "Fuzzy Modelling using Butterfly Optimization Algorithm for Phishing Detection," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 1, pp. 355-360, 2020.

[46] S. Shahbeig, M. Sadegh, and A. Rahideh, "A fuzzy multi-objective hybrid TLBO – PSO approach to select the associated genes with breast cancer," *Signal Processing*, vol. 131, pp. 58-65, 2017, doi: 10.1016/j.sigpro.2016.07.035.

[47] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A New Heuristic Optimization Algorithm: Harmony Search," *Simulation*, vol. 76, no. 2, pp. 60-68, 2001, doi: 10.1177/003754970107600201.

[48] K. Pandiarajan and C. K. Babulal, "Fuzzy harmony search algorithm based optimal power flow for power system security enhancement," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 72-79, 2016, doi: 10.1016/j.ijepes.2015.11.053.

[49] Y. Zhao, J. Xu, and J. Wu, "A New Method for Bad Data Identification of Oilfield System Based on Enhanced Gravitational Search-Fuzzy C-Means Algorithm," *IEEE Trans. Ind. Informatics*, vol. 15, no. 11, pp. 5963-5970, 2019, doi: 10.1109/TII.2019.2935749.

[50] R. Li, S. Su, M. Fan, H. Song, and W. Wang, "A Fuzzy Multi-objective Operation Strategy of Energy Storage System based on Gravity Search Algorithm," *2019 IEEE PES Innov. Smart Grid Technol. Asia*, pp. 480-485, 2019, doi: 10.1109/ISGT-Asia.2019.8881135.

[51] S. Priya, S. Selvakumar, and R. L. Velusamy, "Gravitational Search Based Feature Selection for Enhanced Phishing Websites Detection," *2nd Int. Conf. Innov. Mech. Ind. Appl. ICIMIA 2020 - Conf. Proc.*, no. June, 2020, pp. 453-458, doi: 10.1109/ICIMIA48430.2020.9074837.

[52] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection," *Expert Syst. Appl.*, vol. 116, pp. 147-160, 2019, doi: 10.1016/j.eswa.2018.08.051.

BIOGRAPHIES OF AUTHORS

Noor Syahirah Nordin received the B. Sc. degree in computer science from Universiti Malaysia Pahang (UMP) in 2019. She is currently pursuing the Master of Science in Soft Computing and Intelligent System at UMP.

Mohd Arfian Ismail is a Senior Lecturer at the Faculty of Computing in University Malaysia Pahang, He received B. Sc, M. Sc. and PhD degree in Computer Science from Universiti Teknologi Malaysia (UTM) in 2008, 2011 and 2016, respectively. His current research interests are in the areas of machine learning methods and soft-computing method.

Tole Sutikno is a Lecturer in Electrical Engineering Department at the Universitad Ahmad Dahlan (UAD), Yogyakarta, Indonesia. He received his B.Eng., M.Eng. and Ph.D. degrees in Electrical Engineering from Universitas Diponegoro, Universitas Gadjah Mada and Universiti Teknologi Malaysia, in 1999, 2004 and 2016, respectively. He has been an Associate Professor in UAD, Yogyakarta, Indonesia since 2008. He is currently an Editor-in-Chief of the TELKOMNIKA and the Head of the Embedded Systems and Power Electronics Research Group. His research interests include the field of digital design, industrial applications, industrial electronics, industrial informatics, power electronics, motor drives, renewable energy, FPGA applications, embedded system, artificial intelligence, intelligent control and digital library.

Shahreen Kasim is currently an associate professor in Department of Security Information and Web Technology, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia. Her areas of interest include bioinformatics, soft computing, data mining, web and mobile application.

Rohayanti Hassan is a Senior Lecturer at School of Computing, Universiti Teknologi Malaysia. She has been a lecturer at Software Engineering Department for more than ten years. Her research interests focus on bioinformatics artificial intelligence, software testing and requirement engineering.

Zalmiyah Zakaria is a Senior Lecturer at School of Computing, Universiti Teknologi Malaysia. She received B.Sc., M.Sc. and PhD degrees in Computer Science from Universiti Teknologi Malaysia, in 1999, 2001 and 2009, respectively. Her research interests focus on timetabling/scheduling and machine learning.

Mohd Saberi Mohamad is a Professor of Bioinformatics and Artificial Intelligence at Department of Genetics and Genomics, College of Medicine and Health Sciences, UAEU. His research areas are Artificial Intelligence, Bioinformatics, Data Science, and Computational Biology.