© Universiti Tun Hussein Onn Malaysia Publisher's Office

JSCDM

Journal of Soft Computing and Data Mining

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/jscdm
e-ISSN : 2716-621X

Wrapper Feature Selection Approach Based on Binary Firefly Algorithm for Spam E-mail Filtering

P. D. Veysel Aslantaş[1], D. Ahmet Nusret Toprak[2], Farah Hatem Khorsheed[3], Bashar Ahmed Khalaf[4*]

[1]Department of Computer Engineering,
Erciyes University, Kayseri, 38010, TURKEY

[2]Department of Computer Engineering,
Erciyes University, Kayseri, 38010, TURKEY

[3]Department of Computer Engineering,
Bagdad University, Bagdad, 10001, IRAQ

[4]College of Basic Education,
University of Diyala, 32001, Diyala, IRAQ

*Corresponding Author

DOI: https://doi.org/10.30880/jscdm.2020.01.02.005
Received 11 October 2020; Accepted 22 November 2020; Available online 15 December 2020

Abstract: The current challenges experienced in spam e-mail detection systems is directly associated with the low accuracy of spam e-mail classification and high dimensionality in feature selection processes. However, Feature selection (FS) as a worldwide optimization approach in machine learning (ML) decreases data redundancy and creates a set of accurate and acceptable results. In this paper, a Firefly algorithm-based FS algorithm is proposed for decreasing the dimensionality of features and enhance the accuracy of classifying spam e-mails. The features are represented in a binary form for each firefly; in other words, the features are converted to binary using a sigmoid function. The proposed Binary Firefly Algorithm (BFA) explores the space of the best feature subsets, and the selection of a feature is based on a fitness function which is dependent on the achieved accuracy using Naïve Bayesian Classifier (NBC). The performance of the classifier and the dimension of the selected feature vector as a classifier input are considered when evaluating the performance of the BFA using SpamBase dataset. However, based on the obtained results it is observed that the proposed approach achieved good results with accuracy of 95.14%.

Keywords: Feature selection, firefly algorithm, e-mail spam filtering, classification, Naïve Bayesian

1. Introduction

Globally, e-mails are considered a reliable and best communication channel but recently, this technology has been a major target for attacks. Spam e-mails or junk emails form a large chunk of this attack as they are delivered by different protocols such as simple mail transfer protocol [1], [2]. Being sent in high numbers, these emails occupy a large portion of bandwidth resources when using network resources. They can also deprive users of using network resources as they

*Corresponding author: basharalzubaidy60@gmail.com

tend to block or leverage the available server storage space that is meant for legal users. Similarly, spam e-mails result in a waste of valuable communication time and effort. Consequently, spam e-mails can also be a source of threat to government establishments [3] [4]. Generally, spam e-mail detection is dependent on the appropriate classification of e-mails into spam and non-spam categories.

Most of the recent spam detection frameworks are based on ML techniques for spam e-mails classification [5] [6]. However, a major problem that threatens e-mail classification is the selection of the classifiers' optimal input feature subsets which is to be done through an FS process. Meanwhile, the problem of high data dimensionality which is related to the FS process usually hampers the performance of most classifiers such as the Artificial Neural Network (ANN), Support Vector Machine (SVM), and NBC [7, 8, 9], [10]–[12]. It is assumed that high data dimensionality can be prevented by limiting the feature space and reducing a large number of features in the message. However, it is ideal to identify features concerning the concept of the document or concerning the problems encountered by the document. The accuracy of classification can be affected by irrelevant features. It can also affect the required time to train a classifier, the feature-related cost, and the number of instances required for learning [13], [14].

In recent times, the swarm evolutionary methods such as Ant Colony [15]–[17], Genetic Algorithm [18]–[20], Artificial Bee Colony[21], [22] Particle Swarm Optimization [23], [24] and Harmony Search Algorithm have been employed to solve the FS problems [25], [26]. The Firefly Algorithm (FA) is a swarm-based metaheuristic developed by Yang [27], [28], and has attracted much research attention due to its potential for solving optimization problems [29]. It employs for handling several case studies and problems, such as FS [30]–[33], prediction problems [34], [35], forecasting [36]–[38], scheduling [39], [40], and image processing [41], [42].

In this paper, a wrapper feature selection approach based on a binary firefly algorithm (BFA) is proposed. The BFA selects the best subset features in the Spam base dataset to enhance the filtering rate or classification accuracy of junk emails. The rest of this paper is structured as follows: Section 2 describes the standard FA and NBC, while section 3 explains the proposed algorithm. Section 4 illustrates the experimental results. Finally, the last section provides the conclusion of the study.

2. Firefly Algorithm (FA)

The FA mimics the mating system and data through flashing lights. In this section, the behavior of fireflies, binary fireflies, artificial FA, and FA in some prominent places was discussed.

2.1 The Behavior of Fireflies

There are more than 2,000 species of fireflies worldwide and the vast majority of fireflies send short and pattern flames [27]. The main source of this debt is the interest of accomplices in mating e.g. correspondence, potential, and a component of mechanism. Two factors improve the visibility of most fireflies only at a limited distance [27], only away from the bat, and the brightness of a source at a certain distance corresponds to the law of the opposite square, which suggests that the power of light with expansion decays somewhere$I \propto \frac{1}{r^2}$. The next factor is the assimilation of light, which is recognized around it, which decreases its force as the separation increases.

2.2 Artificial Fireflies

Ordinarily, three idealized rules described the behavior of fireflies as formulated by Yang [27]. These rules are as follows:

- Regardless of gender, all fireflies have only single-sex and can be contracted.
- Firefly attractiveness compared to its brilliance; in this way, prouder fireflies usually attract smaller ones. Attractive quality falls somewhere in the middle of luminosity. Without a bright fire flight, different fire flights move randomly.
- The luminosity of a firefly is determined by the objective function landscape. Then, Brightness/luminosity is directly related to the value of the objective function for the maximization problem. The attraction of a FA i towards extra one with more strength j as Eq. 1:

$$X_i = \beta_0 e^{-\gamma r^2}(X_j - X_i) + \alpha(Rand \ \frac{1}{2}) \quad (1)$$

Where *i* = enchantment, and j is a random coincidence; α = randomization limit, Rand = any number selected for a single transmission in [0; 1]. The articulation range (rand- 0: 5) then starts at [-0,5,0,5] to make room for both positive and negative changes. B is usually 1 and α [0; 1]. α is a disturbance that can affect light transmission. This edge can be selected in a fake fire flight so that the layout can be changed and it now offers an additional layout. The randomization limit can be drawn in the same way for a typical recipe with a mean of 0 and a variation of 1; N (0; 1) to calculate the degree of climate confusion. γ is a seductive variety and its value is important to ensure fusion rate and FA behavior. In many applications, this has changed from 0.01 to 100. Separation of f_i, f_j; means as f_{ij}; and characterized by condition 2.

$$f_{ij} = \left\| X_i - X_j \right\| \tag{2}$$

Where X_i and X_j are the positions of firefly i and j.

Note: In the refreshing state $\beta_0 e_{ij}^{-\gamma r^2}$, an attractive coefficient is used to assume an accident due to the separation of light due to separation according to administrative rules. Besides, the effects of residues and climate on brightness can be specified using an irregular expression of the condition. In this way, the behavior of the firefighter can be determined in the FA pseudocode (see Fig. 1). The characteristics of the FA may be indicated in the following points:

- FA is a versatile skill that supports the benefits of population growth.
- The FA can deal effectively with multi-model problems without many steps, as it allows the population to be divided, with a gradual vision of each leaflet being limited to give them subdivisions in the study area.
- The collection rate of the calculation can be increased by setting random and enchantment limits for the FA throughout the stress cycle.

```
Firefly Algorithm
Input: N = Number of Fireflies
       N Itr = number of iterations for optimization
       β0 Attractiveness parameter
       α Randomization parameter
Output: Optimum FA position
        Initialize a population of n FA random
        find the best solution based on fitness
While Stop Criteria not met do
|      For each Firefly i do
|      |    For each Firefly j do
|      |    |    If Firefly j is better than Firefly I then
|      |    |    |  Move Firefly I towards Firefly j using Equ.1
|      |    |    |        Update Firefly i Intensity
|      |    |    End
|      |    |        Evaluation the position of individual Firefly
|      |    End
|     End
Loop
```

Fig. 1 - The pseudo-code of standard FA

3. The Proposed Binary Firefly Algorithm

Focusing on calculations has evolved because there is a need to find better subgroups of projects that ensure better performance changes. In the proposed model, all fireflies are presented in parallel, categorized, unlike in the traditional risk-prevention model, where fireflies are set up arbitrarily selected highlights. The proposed BFA involves four (4) important advances - promotion, welfare work, inclusive quality counting, and status updates. These tools are explained in more detail in the accompanying subsections.

3.1 BFA Initialization

All fireflies in the search space start in this progression by an irregular number in the interval [0,1]. These irregular numbers tell us about the situation of research. The condition of each firefly is determined by Eq. 3,

$$X = (UB - LB) * Rand(0,1) + LB \tag{3}$$

Where UB is upper bound [1.0], LB = lower bound [0.0], and Rand () =function represent a logistic chaotic map. However, it helps the firefly algorithm to start from positions better than the randomized by a uniform distribution which is given in Eq. 4: -

$$X_{i+1} = \mu X_i (1 - X_i) \quad (4)$$

Where X_i=initial value, X_{i+1}=next value, and μ=control parameter 'mutation'

3.2 Calculation of the fitness function (FF)

In the proposed algorithm, FF must limit the rate of error in the approval agreements for approval, as shown in condition 5, but increase the number of essential or unselected highlights. The FF algorithm was determined by a classifier. Here NBC was used to determine the accuracy of the group.

$$Error = 100 - A \quad (5)$$

where the rating = A = accuracy rate, in other words, 5 times the cross-validation error rate after training NBC. Eq. 6 is used to calculate the intensity of each firefly based on the error value.

$$I(F_i) = \frac{1}{1 + Error^2} \quad (6)$$

3.3 FA Attractiveness Calculation

The Attractive FA has been calculated by several formulas to calculate the level of attractiveness β of each Firefly, Equ. 7 was deployed:

$$\beta(\mathrm{r}) = \beta_0 \times \mathrm{e}^{-\gamma \mathrm{r}^2} \quad (7)$$

where r = the distance between 2 fireflies (calculated using Equ. 8), β_0 = the attractiveness of a firefly at the initial case (r = 0).

$$\mathbf{r_{ij}} = \left| \mathbf{x_i} - \mathbf{x_j} \right| \quad (8)$$

where X represents a real positional value.

Furthermore, the distance between both fireflies can be calculated by using a method of hamming distance. In FA, the distance is exemplified by the change between the two Fireflies binary strings. Also, the employed of FA will develop the ability to work better with binary features and make it working with continuous values.

3.4 Improving FA Position

In the swarm, each firefly is attracted to a brighter Firefly. In the algorithm, the position of the brighter Firefly is updated using Equ. 9.

$$X_i = X_i + \beta \times (X_j - X_i) + \alpha \times (Rand - \frac{1}{2}) \quad (9)$$

Where x_i in the first part of the relation = current position of the best firefly, while the second part of the relation expresses the attractiveness between position F_i and F_j. Gr represents the information gain ratio values for all properties previously calculated in the first step [43]. The third stage of the relative expresses the randomization through α, anywhere α ∈ (0,1). This chance is decreased by another constant rate δ, where δ ∈ (0.95, 0.97) such that at the final optimization stage, the value of α will increase, as in Eq. (10).

$$\alpha = \alpha \times \delta \quad (10)$$

4. Results And Discussions

Some evaluation metrics were used to evaluate the performance of the proposed BFA based on the selected Spam Base dataset. The simplest evaluation measure is filtering accuracy, which is a measure of the percentage of messages that are correctly classified. The accuracy (determined using Eq. 11) is the percentage of emails that are correctly identified as spam and not spam.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

Another evaluation measure is recall and precision. The recall is the percentage of spam emails that are blocked, while precision is the percentage of correct messages that are marked as spam. Both recall and precision are calculated using Eq.s 12 and 13

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

The SPAMBASE dataset was accessed from the UCI machine learning repository [44]. It was created by Mark Hopkins and Co. as a dataset containing 4601 email messages and 58 attributes. The non-spam emails in this dataset were collected from personal e-mails, field works, and single e-mail accounts. The set of emails contained in this dataset are suitable for testing spam filtering systems. In the SPAMBASE, each instance is made up of 58 attributes and most of these attributes are the frequency of a given character in the email which corresponds to the instance.

There are two main parameters of the BFA; the first is the swarm size (SS) which indicates the number of fireflies in the swarm, and the second is the MaxITr which indicates the number of iterations. The dataset was divided into two parts, 70% for training and 30% for testing. The BFA was executed for 20 runtimes using different swarm sizes and several iterations to compare its performance in finding the best subset of features with higher accuracy with (XX) algorithms. All the experiments were carried out on a standalone PC with 4 GB of RAM and 2.2GHz core i5 of CPU. The algorithm was written and executed using C#.net 5.0 programming language. In this study, four swarm sizes (10, 20, 30, 40, and 50) were used, and each swarm size was tested with different numbers of iterations (100, 250, 300, and 500). Table (1) shows the results of these experiments.

Table 1 shows that the accuracy of the BFA has continued to increase with the extension of the crowd, suggesting an impact on the size of the group of facts. Besides, the range of actions leads to a further impact on the measurement system. Therefore, it can be argued that the small size and number of circles have a positive effect on the reliability of the BFA. At the end of the day, the calculation will be moderate if it will be larger and more complex of cycles. The benefits of this restriction are shown in Fig. 2.

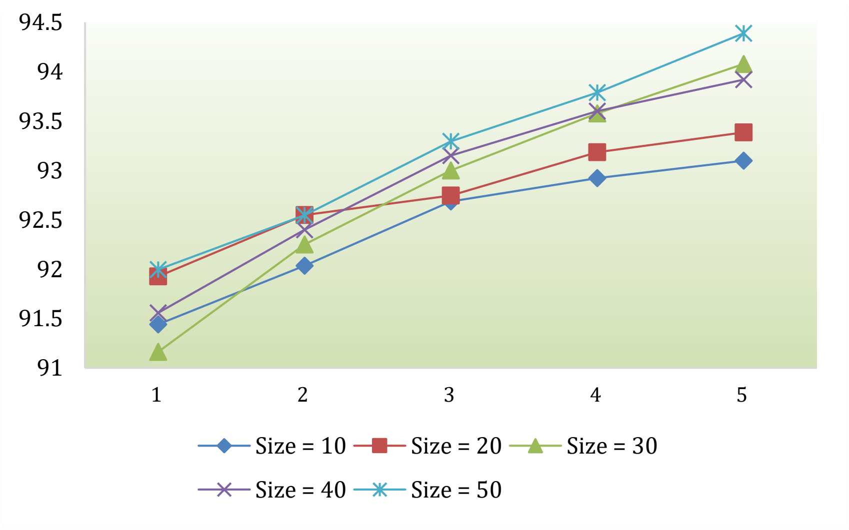

Fig. 2 - The effect of swarm size and number of iteration on the accuracy of BFA

Table 1 - The results of the proposed BFA

MaxItr	Swarm Size	Best Accuracy	Worst Accuracy	Average Accuracy	Average Features
	10	92.91	90.22	91.441	35
	20	93.06	90.99	92.548	33.6
100	30	94.14	91.48	92.746	31
	40	94.15	91.95	92.922	29.8
	50	94.251	92.33	93.099	28
	10	93.16	90.03	91.926	33.4
	20	93.39	92.063	92.053	31.9
200	30	93.69	92.736	92.686	30.6
	40	94.15	93.185	93.185	28.4
	50	94.25	93.368	93.386	25.7
	10	92.72	90.22	91.163	34.8
	20	93.77	90.8	92.285	33.4
250	30	94.34	92.24	93	29.2
	40	94.7	92.74	93.578	27.5
	50	94.9	93.57	94.078	25
	10	93.58	90.03	91.557	34.4
	20	93.39	90.9	92.25	29.2
300	30	93.76	92.95	93.294	27.3
	40	94.2	93.19	93.6	26.5
	50	94.91	93.29	93.919	23.2
	10	93.29	90.7	91.997	30.5
	20	93.29	91.76	92.547	29.7
500	30	93.39	92.08	93.15	26.6
	40	94.81	93.33	93.789	23.5
	50	95.14	93.62	94.389	21.6

Table 2 shows the comparison between the accuracy of the proposed BFA and three standard classification models, support vector machine (SVM), and K-nearest neighbor (KNN), naïve Bayesian classifier (NBC). The table below illustrates the impact of the proposed feature selection algorithm on the classification accuracy, meaning that, the stander classifiers used all features in the dataset (i.e. 57 features), while the proposed algorithm selected a subset of 21 features which enhances the classification accuracy to 95.14.

However, based on the obtained results it is observed that the proposed system achieved the highest results during comparison with the related work. It is worth to mention that the standard NBC attained 79.6 with all features, while it attained 95.14 when the features selection algorithm or firefly algorithm is applied. Table 2 shows the comparison of the proposed system with the related work.

Table 2 - The Comparison of the proposed FA and three standard models

Model	Acc. Rate	Err. Rate	No. of Features
SVM	90.42	9.58	57
KNN	89.52	10.48	57
NBC	79.6	20.4	57
Proposed System	95.14	4.86	21

Moreover, the proposed model performance has been compared with the most related feature selection algorithm which are ACO-SVM, ABC-SVM, GA-NBC, and ACO-NBC as shown in the below table:

Table 3 - The comparison of the proposed approach with the three standard models

No	Algorithm	Acc. Rate	Err. Rate	Ref
1	ACO-SVM	81.25	19.75	[45]
2	ABC-SVM	67.9	32.1	[46]
3	GA-NBC	77	23	[47]
4	ACO-NBC	84	16	[47]
5	Proposed System	95.14	4.86	-

5. Conclusion

In this study, the Firefly algorithm was selected for the most appropriate accents that would improve NBC's accuracy and predictability. Firefly's algorithm was based on a cluttered strategic guide before doubling its position using sigmoidal power. NBC was used in the proposed algorithm as an assessment of order well-being. Overall, NBC achieved low accuracy (79.5%), in contrast to the usual KNN or SVM. The Firefly algorithm has greatly improved NBC's accuracy by at least more than 90%, indicating a better presentation of the proposed algorithm that does not meet the standard SVM and KNN. Tests have shown an amount that would affect the presentation of the Firefly algorithm. The accuracy of the classifier has been extended from set to set. Also, weight had a negligible effect on the display of the classification. This ratio showed that the proposed algorithm outperformed other comparable algorithms such as ACO, ABC, and GA.

Acknowledgement

This work was supported by the department of Computer, College of Basic Education, University of Diyala, 32001, Diyala, Iraq.

References

[1] DeBarr, D., & Wechsler, H. (2012). Spam detection using random boost. Pattern Recognition Letters, 33(10), 1237-1244.

[2] Wu, Q., Wu, S., & Liu, J. (2010). Hybrid model based on SVM with Gaussian loss function and adaptive Gaussian PSO. Engineering Applications of Artificial Intelligence, 23(4), 487-494.

[3] Lee, S. M., Kim, D. S., Kim, J. H., & Park, J. S. (2010, February). Spam detection using feature selection and parameters optimization. In 2010 International Conference on Complex, Intelligent and Software Intensive Systems (pp. 883-888). IEEE.

[4] Jindal, N., & Liu, B. (2007, October). Analyzing and detecting review spam. In Seventh IEEE International Conference on Data Mining (ICDM 2007) (pp. 547-552). IEEE.

[5] Hall, M. A. (1999). Correlation-based feature selection for machine learning.

[6] Chang, M., & Poon, C. K. (2009). Using phrases as features in email classification. Journal of Systems and Software, 82(6), 1036-1045.

[7] Zhao, M., Fu, C., Ji, L., Tang, K., & Zhou, M. (2011). Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. Expert Systems with Applications, 38(5), 5197-5204.

[8] Liu, H., Shi, X., Guo, D., & Zhao, Z. (2015). Feature selection combined with neural network structure optimization for HIV-1 protease cleavage site prediction. BioMed research international.

[9] Li, S., Wang, P., & Goel, L. (2015). Wind power forecasting using neural network ensembles with feature selection. IEEE Transactions on sustainable energy, 6(4), 1447-1456.

[10] Chen, J., Huang, H., Tian, S., & Qu, Y. (2009). Feature selection for text classification with Naïve Bayes. Expert Systems with Applications, 36(3), 5432-5435.

[11] Fadel, H., Hameed, R. S., Hasoon, J. N., & Mostafa, S. A. (2020). A Light-weight ESalsa20 Ciphering based on 1D Logistic and Chebyshev Chaotic Maps. *Solid State Technology*, *63*(1), 1078-1093.

[12] Feng, G., Guo, J., Jing, B. Y., & Sun, T. (2015). Feature subset selection using naive Bayes for text classification. Pattern Recognition Letters, 65, 109-115.
[13] Unler, A., Murat, A., & Chinnam, R. B. (2011). mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification. Information Sciences, 181(20), 4625-4641.
[14] Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence, 27(8), 1226-1238.
[15] Moradi, P., & Rostami, M. (2015). Integration of graph clustering with ant colony optimization for feature selection. Knowledge-Based Systems, 84, 144-161.
[16] Paniri, M., Dowlatshahi, M. B., & Nezamabadi-pour, H. (2020). MLACO: A multi-label feature selection algorithm based on ant colony optimization. Knowledge-Based Systems, 192, 105285.
[17] M. Ghosh, R. Guha, R. Sarkar, and A. Abraham, "A wrapper-filter feature selection technique based on ant colony optimization," Neural Comput. Appl., pp. 1–19, 2019.
[18] Ghosh, M., Guha, R., Sarkar, R., & Abraham, A. (2019). A wrapper-filter feature selection technique based on ant colony optimization. Neural Computing and Applications, 1-19.
[19] Kabir, M. M., Shahjahan, M., & Murase, K. (2011). A new local search based hybrid genetic algorithm for feature selection. Neurocomputing, 74(17), 2914-2928.
[20] Lin, C. H., Chen, H. Y., & Wu, Y. S. (2014). Study of image retrieval and classification based on adaptive features using genetic algorithm feature selection. Expert Systems with Applications, 41(15), 6611-6621.
[21] Schiezaro, M., & Pedrini, H. (2013). Data feature selection based on Artificial Bee Colony algorithm. EURASIP Journal on Image and Video Processing, 2013(1), 47.
[22] Agrawal, V., & Chandra, S. (2015, August). Feature selection using Artificial Bee Colony algorithm for medical image classification. In 2015 Eighth International Conference on Contemporary Computing (IC3) (pp. 171-176). IEEE.
[23] Xue, B., Zhang, M., & Browne, W. N. (2014). Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. Applied soft computing, 18, 261-276.
[24] Xue, B., Zhang, M., & Browne, W. N. (2012). Particle swarm optimization for feature selection in classification: A multi-objective approach. IEEE transactions on cybernetics, 43(6), 1656-1671.
[25] Ramos, C. C., Souza, A. N., Chiachia, G., Falcão, A. X., & Papa, J. P. (2011). A novel algorithm for feature selection using harmony search and its application for non-technical losses detection. Computers & Electrical Engineering, 37(6), 886-894.
[26] Inbarani, H. H., Bagyamathi, M., & Azar, A. T. (2015). A novel hybrid feature selection method based on rough set and improved harmony search. Neural Computing and Applications, 26(8), 1859-1880.
[27] Yang, X. S. (2009, October). Firefly algorithms for multimodal optimization. In International symposium on stochastic algorithms (pp. 169-178). Springer, Berlin, Heidelberg.
[28] Yang, X. S. (2010). Firefly algorithm, Levy flights and global optimization. In Research and development in intelligent systems XXVI (pp. 209-218). Springer, London.
[29] Yang, X. S., & He, X. (2013). Firefly algorithm: recent advances and applications. International journal of swarm intelligence, 1(1), 36-50.
[30] Anbu, M., & Mala, G. A. (2019). Feature selection using firefly algorithm in software defect prediction. Cluster Computing, 22(5), 10925-10934.
[31] Al-Jawhar, Y. A., Ramli, K. N., Taher, M. A., Shah, N. S. M., Mostafa, S. A., & Khalaf, B. A. Improving PAPR performance of filtered OFDM for 5G communications using PTS. *ETRI Journal*.
[32] Abatunde, O.S., Ahmad, A.R., Mostafa, S.A., khalaf, B. A, Fadel, A.H., Shamala, P. (2020). A smart network intrusion detection system based on network data analyzer and support vector machine. International Journal of Emerging Trends in Engineering Research, 8(1 Special Issue 1), pp. 213-220
[33] Banati, H., & Bajaj, M. (2011). Fire fly based feature selection approach. International Journal of Computer Science Issues (IJCSI), 8(4), 473.
[34] Obaid, O. I., Mohammed, M. A., Ghani, M. K. A., Mostafa, A., & Taha, F. (2018). Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer. *International Journal of Engineering & Technology*, *7*(4.36), 160-166.
[35] Ghorbani, M. A., Deo, R. C., Yaseen, Z. M., Kashani, M. H., & Mohammadi, B. (2018). Pan evaporation prediction using a hybrid multilayer perceptron-firefly algorithm (MLP-FFA) model: case study in North Iran. Theoretical and applied climatology, 133(3-4), 1119-1131.
[36] Yaseen, Z. M., Ghareb, M. I., Ebtehaj, I., Bonakdari, H., Siddique, R., Heddam, S., ... & Deo, R. (2018). Rainfall pattern forecasting using novel hybrid intelligent model based ANFIS-FFA. Water resources management, 32(1), 105-122.

[37] Khalaf, B. A., Mostafa, S. A., Mustapha, A., Mohammed, M. A., & Abduallah, W. M. (2019). Comprehensive review of artificial intelligence and statistical approaches in distributed denial of service attack and defense methods. *IEEE Access*, *7*, 51691-51713.
[38] Mehr, A. D., Nourani, V., Khosrowshahi, V. K., & Ghorbani, M. A. (2019). A hybrid support vector regression–firefly model for monthly rainfall forecasting. International Journal of Environmental Science and Technology, 16(1), 335-346.
[39] Kashikolaei, S. M. G., Hosseinabadi, A. A. R., Saemi, B., Shareh, M. B., Sangaiah, A. K., & Bian, G. B. (2020). An enhancement of task scheduling in cloud computing based on imperialist competitive algorithm and firefly algorithm. The Journal of Supercomputing, 76(8), 6302-6329.
[40] Karthikeyan, S., Asokan, P., Nickolas, S., & Page, T. (2015). A hybrid discrete firefly algorithm for solving multi-objective flexible job shop scheduling problems. International Journal of Bio-Inspired Computation, 7(6), 386-401.
[41] Dong, H., He, M., & Qiu, M. (2015, September). Optimized gray-scale image watermarking algorithm based on DWT-DCT-SVD and chaotic firefly algorithm. In 2015 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (pp. 310-313). IEEE.
[42] Hrosik, R. C., Tuba, E., Dolicanin, E., Jovanovic, R., & Tuba, M. (2019). Brain image segmentation based on firefly algorithm combined with k-means clustering. Stud. Inform. Control, 28, 167-176.
[43] Ismael, H. A., Abbas, J. M., Mostafa, S. A., & Fadel, A. H. (2020). An enhanced fireworks algorithm to generate prime key for multiple users in fingerprinting domain. *Bulletin of Electrical Engineering and Informatics*, *10*(1), 337-343.
[44] UCI Machine Learning Repository, “Spambase Dataset,” University of California, School of Information and Computer Science. [Online]. Available: http://archive.ics.uci.edu/ml/datasets.