An Enhanced Particle Swarm Optimization algorithm for E-mail Spam Filtering

Article *in* AUS · August 2019

DOI: 10.4206/aus.2019.n26.2.31

5 authors, including:

Khattab M. Ali Alheeti
University of Anbar

Saif Saad
University of Anbar

♦ Recepción/ 27 junio 2019
♦ Aceptación/ 25 agosto 2019

An Enhanced Particle Swarm Optimization algorithm for E-mail Spam Filtering

Un algoritmo mejorado de optimización de enjambre de partículas para el filtrado de correo electrónico no deseado

Hadeel M.Saleh
College of Education for Women, University of Anbar, Ramadi, Iraq.
E-mail: Hadealms89@gmail.com

Khattab M. Ali Alheeti
Computer Networking Systems Department College of Computer Science and Information Technology, University of Anbar
E-mail: co.khattab.alheeti@uonabr.edu.iq

Saif Saad Hameed
Computer Networking Systems Department College of Computer Science and Information Technology, University of Anbar
E-mail: dovewhite89@gmail.com

Omer. Y. Assaf
Ministry of education
E-mail: ojonaid@yahoo.com

Noor M. Jassam
University of Anbar
E-mail: Noor82noor2017@gmail.com

ABSTRACT/ The current challenges experienced in spam email detection systems is directly associated with the low accuracy of spam email classification and high dimensionality in feature selection processes. However, Feature selection (FS) as a global optimization approach in machine learning decreases data redundancy and creates a set of accurate and acceptable results. In this paper, a particle swarm optimization (PSO) algorithm is enhanced by using a logistic chaotic map for decreasing the dimensionality of features and enhance the accuracy of classifying spam emails. The features are represented in a binary from for each particle; in other words, the features are converted to binary using a sigmoid function. The selection of the features is based on a fitness function which is dependent on the achieved accuracy using Support Vector Machine (SVM). The performance of the classifier and the dimension of the selected features vector as a classifier input are considered when evaluating the performance of the Chaotic Binary PSO (CBPSO) using SpamBase dataset. The outcome of the experiments showed the BPSO to achieve good FS results even with a small set of selected features.
Keywords: Particle Swarm Optimization; feature selection; e-mail spam filtering; SVM
RESUMEN/ Los desafíos actuales experimentados en los sistemas de detección de correo no deseado están directamente asociados con la baja precisión de la clasificación del correo no deseado y la alta dimensionalidad en los procesos de selección de funciones. Sin embargo, la selección de características (FS) como un enfoque de optimización global en el aprendizaje automático disminuye la redundancia de datos y crea un conjunto de resultados precisos y aceptables. En este documento, se mejora un algoritmo de optimización de enjambre de partículas (PSO) mediante el uso de un mapa caótico logístico para disminuir la dimensionalidad de las funciones y mejorar la precisión de la clasificación de correos electrónicos no deseados. Las características se representan en un binario de para cada partícula; en otras palabras, las características se convierten en binarias usando una función sigmoidea. La selección de las características se basa en una función de condición física que depende de la precisión lograda con Support Vector Machine (SVM). El rendimiento del clasificador y la dimensión del vector de características seleccionado como entrada del clasificador se consideran al evaluar el rendimiento del Chaotic Binary PSO (CBPSO) utilizando el conjunto de datos SpamBase. El resultado de los experimentos mostró que el BPSO logró buenos resultados de FS incluso con un pequeño conjunto de características seleccionadas. Palabras clave: Optimización de enjambre de partículas; selección de características; filtrado de correo electrónico no deseado; SVM.

1. Introduction

Globally, e-mails are considered as a reliable and best communication channel but recently, this technology has been a major target for attacks. Spam emails or junk emails form a large chunk of this attack as they are delivered by different protocols such as simple mail transfer protocol (SMTP)[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 tend to block or leverage the available server storage space that is meant for legal users. Similarly, spam emails result in the waste of valuable communication time and effort. Consequently, spam emails can also be a source of threat to government establishments [3] [4]. Generally, spam email detection is dependent on the appropriate classification of emails into spam and non-spam categories.

Most of the recent spam detection frameworks are based on ML techniques for spam emails classification [5] [6]. However, a major problem that threatens email classification is the selection of the classifiers' optimal input feature subsets which is to be done through a 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 the large number of features in the message. However, it is ideal to identify features with respect to the concept of the document or with respect to 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].

Recently, the swarm-based and evolutionary methods such as Ant Colony Optimization (ACO)[15]–[17], Genetic Algorithm (GA) [18]–[20], Artificial Bee Colony (ABC)[21], [22] Particle Swarm Optimization (PSO)[23], [24] and Harmony Search Algorithm (HSA) have been used to handle the problems of FS [25], [26]. The particle swarm optimization (PSO) algorithm is one of the researched nature-inspired swarm intelligence framework [23], [27], [28]. The PSO is inspired by hunting or living styles of birds and fish. It has found application in the handling of complex optimization tasks. It was introduced by [27] and since its introduction, it has undergone several modifications which result in several PSO variants which are aimed at finding a better way of handling specific optimization problems. There are four categories of the modifications on the PSO variants: The first category of modification is centered on the parameter settings with emphasis on the inertia weight and acceleration coefficients parameters optimization. In the second variants, the emphasis is on the topology of the neighborhood which expresses the inter-particle connectivity. The third category of modification focus on the learning strategies with an emphasis on teaching and peer learning of the bests and global best positions of the particles. The fourth category of modification is mainly on the hybridization of the PSO variants with other optimization frameworks [29]–[33].

In this paper, a wrapper feature selection approach based on logistic chaotic map and binary particle swarm optimization (BCPSO) is proposed. The BCPSO 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. The Proposed Methodology

2.1 Particle Swarm Optimization

PSO is one of the metaheuristics inspired by natural occurrences. It was inspired by the flocking pattern of birds when searching for food. The flock of birds in the PSO is considered randomly distributed and there is only one piece of food available to them as depicted in Figure 2 where the only piece of food is marked as a dot on the tree. Although this piece of food is placed at a known distance to the birds, this position is not known to the birds. Only the bird that is nearest to the food source can send signal to the distant birds to flock towards the food source. Therefore, each bird in the flock is considered as a particle while the food piece is considered as the optimal value. The distance from each bird to the food piece is represented by the value of the OF; hence, the flocking behaviour of these birds can be considered as a function optimization process. In the flock, the nearest bird to the food piece is represented as X_i (Figure 1) and as such, is considered as the current global best whose distance from the optimal is given as Nbesti [27], [34]. The basic concept of the PSO considered a specified position and velocity for each particle during an active search for the optimal solution to any NP-hard problem. It also conceptualized the fact that the position of each particles can be iteratively updated based on its existing local and global optima. Thus, each particles' position (e.g., particle i) can be updated thus :

$$X_i(t+1) = X_i(t) + V_i(t+1) \qquad (1)$$

where t= current status, $t+1$ = post-update status, X_i(t+1) = velocity of the new particles. Note that the difference in time $\Delta t = (t+1) - 1$ is the time unit while the particles' velocity i is expressed as:

$$V_i(i+t) = \omega V_i(t) + c_i r_i \left(X_i^P - X_i(t)\right) + c_2 r_2 \left(X^G - X_i(t)\right) \qquad (2)$$

where $v_i(t)$ = current particles' velocity, X_i^P = particles' local best position, X^G = swarm level particles' global best position, ω c_1 , and c_2 = constants that weighs the relevance of each velocity component, r_1 and r_2 = random values in the range of [0, 1].

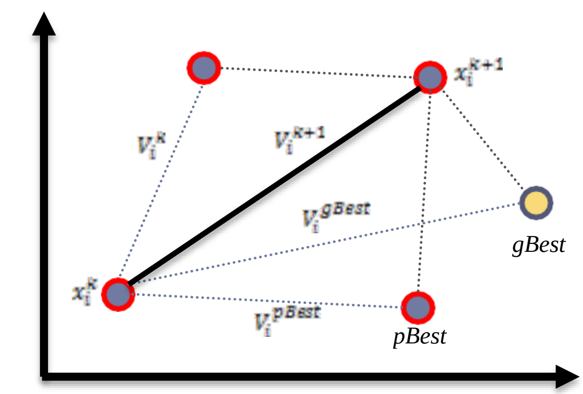

Figure 1. A depiction of the PSO

Despite the several alteration to the PSO framework, there are immobile problems in the PSO that needs to be resolved. These problems include its premature convergence in multimodal functions as it continues to look for the best solutions despite having achieved one, its pace of convergence (can still be stuck in local minima despite previously achieving the best solution),integrated complexity, discontinuity, and multimodality issues that often leads to low quality solutions, solution instability(cannot be acted upon)[32], [33] .

2.2 Chaotic Binary Particle Swarm Optimization (CBPSO)

Feature selection algorithms are mainly developed due to the need for finding better subset features that will offer better performance accuracies. In the proposed algorithm, all the particles in the swarm are where X_i= a real value in range 0 and 1, initialized based on random positions generated using logistic map, then these positions are converted into binary form. Before the algorithm is executed, the data set should be read and normalized. The pseudocode of CBPSO is given in figure 2, the algorithm is comprised of six major steps, as follows:

Step1: Generate the initial position of the particles using the equation of the chaotic logistic map, as follows :

$$X_{i+1} = \mu X_i \, (1 - X_i) \qquad (3)$$

which represents a single dimension of any given problem μ and = the control parameter – or mutation – of logistic map, which is in range 0 and 4.

Step2: Convert all values of Xi into binary form using the sigmoid function, which is given in the following equation :

$$F_i = \begin{cases} 1, & sigmoid\ (X_i) > u\ [0\ ,1] \\ 0, & otherwise \end{cases} \qquad (4)$$

Where: X_i = position of each particle, $sigmoid(X_i)$ = 1 / [1 + e^{-X}], u = uniform distribution, F_i = binary sequence, 1 = chances that a feature will be selected, and 0 = chances that a feature will not be selected.

Step3: Calculate the fitness function for each particle in the swarm based on their binary sequence. The objective function for the proposed case study is the classification accuracy obtained by using a machine learning model, which is Support Vector Machine (SVM). Therefore, the proposed algorithm in this paper is trying to maximize the fitness function; in other words, maximize the classification accuracy of SVM by selecting the most relevant features. The classification accuracy is calculated using the following equation :

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (5)$$

Where: TP represents the true positive, TN represents the True negative, FP represents the false positive, and FP represents the false negative. All these parameters are calculated

after obtaining the confusion matrix of SVM classifier.
Step4: Update the velocity and position for each particle based via equations (2) and (1). The positions after the update should be converted into binary using step 2. The fitness function should be re-calculated using Step 3.

Step5: Update the local best solution ($lbest$), and the global best solution ($gbest$).
Step6: If the stop condition is satisfied, then the algorithm should stops and return the $gbest$. Otherwise, Go to step Step 4

CBPSO Algorithm

```
Read the PSO parameters (          ), MaxIt
Read the SpamBase dataset.
Normalize the features.
Generate the initial positions of the particles via equation 3.
Convert the position into binary via equation 4.
Calculate the objective function by using SVM.
While (t < MaxIt)
    For each particles (p) in the swarm
        Update the velocity of p of via equation 2
        Update the position of p via equation 1
        Update the fitness value of p
    Next
    Update lbest, gbest

Next
```

Figure 2: The pseudocode of CBPSO

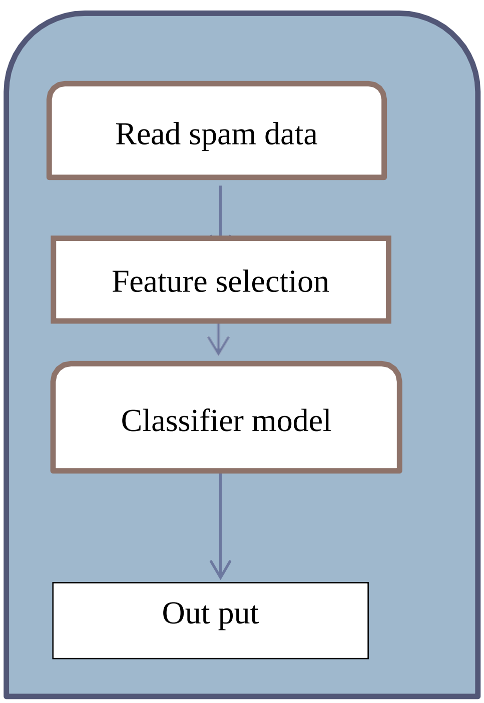

Figure3: The flow chart of CBPSO

3. Results and Discussion

Some evaluation metrics were used to evaluate the performance of the proposed CBPSO based on the selected SpamBase dataset. The simplest evaluation measure is the filtering accuracy, which is a measure of the percentage of messages that are correctly classified [35]. The accuracy (determined using equation 5) is the percentage of emails that are correctly identified as spam and not spam.

The SPAMBASE dataset was accessed from the UCI machine learning repository [36]. It was created by Mark Hopkins and Co. as a dataset containing 4601 email messages and 57 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 57 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 CBPSO; 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 CBPSO was executed for 20 runtimes using different swarm sizes and number of iterations to compare its performance in finding the best subset of features with higher accuracy with 95.14 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,five 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. The results of CBPSO on different no. of iterations and swarm sizes

MaxItr	SS	Best	Worst	Avg. Accuracy	Avg. Feature
100	10	92.91	90.22	91.441	35
	20	93.06	90.99	92.548	33.6
	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
200	10	93.16	90.03	91.926	33.4
	20	93.39	92.063	92.053	31.9
	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
250	10	92.72	90.22	91.163	34.8
	20	93.77	90.8	92.285	33.4
	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
300	10	93.58	90.03	91.557	34.4
	20	93.39	90.9	92.25	29.2
	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
500	10	93.29	90.7	91.997	30.5
	20	93.29	91.76	92.547	29.7
	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

From Table 1, the accuracy of the CBPSO was observed to be gradually increased with an increase in the swarm size, suggesting a possible effect of the swarm size on the accuracy. Additionally, the number of iterations also affected the searching process. Hence, it can be concluded that both swarm size and number of iterations have significant effects on the prediction accuracy of the CBPSO. The effect of these parameters is depicted in Figure 4 .

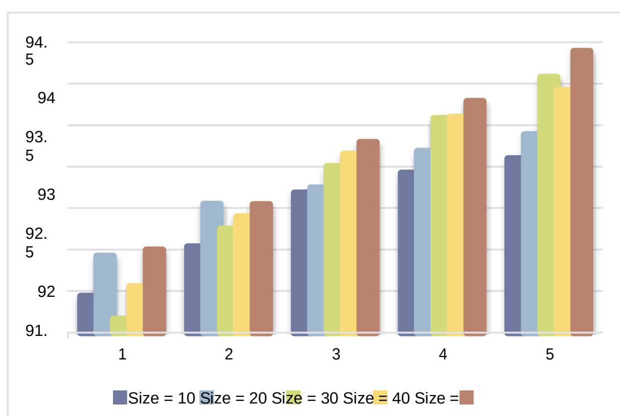

Figure4. The effect of swarm size and number of iteration on the accuracy of CBPSO

Table 2 shows the comparison between the accuracy of the proposed CBPSO and those of three standard classification models (SVM, KNN, and NBC).

Table 2. Comparison of the proposed CBPSO and three standard models

N	Model	Acc. Rate	Err. Rate	No. of Features
1	SVM	90.42	9.58	57
2	KNN	89.52	10.48	57
3	NBC	79.6	20.4	57
4	**Proposed**	**95.14**	**4.86**	**21**

The performance of the proposed CBPSO was further compared to those of 4 other feature selection algorithms (ACO-SVM, ABC-SVM, GA-NBC, and ACO-NBC) as depicted in Table3.

Table 3. Comparison of the proposed algorithm with 3 standard models

N	Algorithm	Acc. Rate	Err. Rate	Ref
1	ACO-SVM	81.25	19.75	[37]
2	ABC-SVM	67.9	32.1	[38]
3	GA-NBC	77	23	[39]
4	ACO-NBC	84	16	[39]
5	**Proposed**	**95.14**	**4.86**	

4. Limitation

This work applies to solve email spam problem using feature selection, this work is limited on this problem, in the future works, we are developing that by using another way, such as apply this work on medical data set like diabetes, heart, cancer diseases or apply that on networks problems another way is implementing this problem on another methodology.

5. Conclusion

In this study, the PSO has been improved and used to select the most relevant features that will enhance the accuracy and prediction performance of the SVM. It was initialised by a chaotic logistic map before converting its positions to binary using the sigmoid function. The SVM was used in the proposed algorithm as a fitness function for evaluating the solutions. Overall, the NBC achieved a low accuracy (79.5%) compared to the standard Knn or SVM. The proposed CBPSO generally enhanced the accuracy of the SVM to a minimum of more than 90%, thereby, suggesting a better performance of the proposed algorithm compared to the standard SVM and Knn. The experiments showed the swarm size to have an influence on the performance of the PSO algorithm. The accuracy of the classifier increased with the number of swarms. Furthermore, the number of iterations was shown to slightly affect the performance of the classifier. The comparison showed that the proposed algorithm outperformed the other benchmarking algorithms such as ACO, ABC, and GA .

References

[1] D. Debarr and H. Wechsler, "Spam detection using Random Boost," *Pattern Recognit. Lett.*, vol. 33, no. 10, pp. 1237–1244, 2012.

[2] Q. Wu, S. Wu, and J. Liu, "Hybrid model based on SVM with Gaussian loss function and adaptive Gaussian PSO," *Eng. Appl. Artif. Intell.*, vol. 23, no. 4, pp. 487–494, 2010.

[3] S. M. Lee, D. S. Kim, J. H. Kim, and J. S. Park, "Spam Detection Using Feature Selection and Parameters Optimization," in *2010 International Conference on Complex, Intelligent and Software Intensive Systems*, 2010, pp. 883–888.

[4] N. Jindal and B. Liu, "Analyzing and detecting review spam," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2007, pp. 547–552.

[5] M. Hall, "Correlation-based Feature Selection for Machine Learning," *Methodology*, vol. 21i195-i20, no. April, pp. 1–5, 1999.

[6] M. Chang and C. K. Poon, "Using phrases as features in email classification," *J. Syst. Softw.*, vol. 82, no. 6, pp. 1036–1045, 2009.

[7] H. Liu, X. Shi, D. Guo, Z. Zhao, and Yimin, "Feature selection combined with neural network structure optimization for HIV-1 protease cleavage site prediction," *Biomed Res. Int.*, vol. 2015, 2015.

[8] S. Li, P. Wang, and L. Goel, "Wind Power Forecasting Using Neural Network Ensembles with Feature Selection," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1447–1456, 2015.

[9] M. Zhao, C. Fu, L. Ji, K. Tang, and M. Zhou, "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes,"

Expert Syst. Appl., vol. 38, no. 5, pp. 5197–5204, 2011.
[10]J. Chen, H. Huang, S. Tian, and Y. Qu, "Feature selection for text classification with Naïve Bayes," *Expert Syst. Appl.*, vol. 36, no. 3 PART 1, pp. 5432–5435, 2009.
[11]M. L. Zhang, J. M. Peña, and V. Robles, "Feature selection for multi-label naive Bayes classification," *Inf. Sci. (Ny).*, vol. 179, no. 19, pp. 3218–3229, 2009.
[12]G. Feng, J. Guo, B.-Y. Jing, and T. Sun, "Feature subset selection using naive Bayes for text classification," *Pattern Recognit. Lett.*, vol. 65, pp. 109–115, 2015.
[13]A. Unler, A. Murat, and R. B. Chinnam, "Mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification," *Inf. Sci. (Ny).*, vol. 181, no. 20, pp. 4625–4641, 2011.
[14]H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, 2005.
[15]P. Moradi and M. Rostami, "Integration of graph clustering with ant colony optimization for feature selection," *Knowledge-Based Syst.*, vol. 84, pp. 144–161, 2015.
[16]A. Al-ani, "Ant Colony Optimization for Feature Subset Selection," vol. 1, no. 4, pp. 999–1002, 2007.
[17]S. M. Vieira, J. M. C. Sousa, and T. A. Runkler, "Two cooperative ant colonies for feature selection using fuzzy models," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 2714–2723, 2010.
[18]C.-H. Cheng, T.-L. Chen, and L.-Y. Wei, "A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting," *Inf. Sci. (Ny).*, vol. 180, no. 9, pp. 1610–1629, 2010.
[19]M. M. Kabir, M. Shahjahan, and K. Murase, "A new local search based hybrid genetic algorithm for feature selection," *Neurocomputing*, vol. 74, no. 17, pp. 2914–2928, 2011.
[20]C.-H. Lin, H.-Y. Chen, and Y.-S. Wu, "Study of image retrieval and classification based on adaptive features using genetic algorithm feature selection," *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6611–6621, 2014.
[21]M. Schiezaro and H. Pedrini, "Data feature selection based on Artificial Bee Colony algorithm," *J. Image Video Process.*, vol. 1, no. 47, pp. 1–8, 2013.
[22]V. Agrawal and S. Chandra, "Feature Selection using Artificial Bee Colony Algorithm for Medical Image Classification," *2015 Eighth Int. Conf. Contemp. Comput.*, vol. 1, pp. 2–7, 2015.
[23]B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms," *Appl. Soft Comput. J.*, vol. 18, pp. 261–276, 2014.
[24]B. Xue, M. Zhang, S. Member, and W. N. Browne, "Particle Swarm Optimization for Feature Selection in Classification : A Multi-Objective Approach," pp. 1–16, 2012.
[25]C. C. O. Ramos, A. N. Souza, G. Chiachia, A. X. Falc??o, and J. P. Papa, "A novel algorithm for feature selection using Harmony Search and its application for non-technical losses detection," *Comput. Electr. Eng.*, vol. 37, no. 6, pp. 886–894, 2011.
[26]H. H. Inbarani, M. Bagyamathi, and A. T. Azar, "A novel hybrid feature selection method based on rough set and improved harmony search," *Neural Comput. Appl.*, vol. 26, no. 8, pp. 1859–1880, 2015.
[27]R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. pp. 39–43, 1995.
[28]J. Kennedy and R. Eberhart, "Particle swarm optimization," *Neural Networks, 1995. Proceedings., IEEE Int. Conf.*, vol. 4, pp. 1942–1948 vol.4, 1995.
[29]İ. B. Aydilek, "A hybrid firefly and particle swarm optimization algorithm for computationally expensive numerical problems," *Appl. Soft Comput. J.*, vol. 66, pp. 232–249, 2018.
[30]K. Premalatha and a M. Natarajan, "Hybrid PSO and GA for Global Maximization," *Int. J. Open Probl. Compt. Math*, vol. 2, no. 4, pp. 597–608, 2009.
[31]D. Chen, J. Chen, H. Jiang, F. Zou, and T. Liu, "An improved PSO algorithm based on particle exploration for function optimization and the modeling of chaotic

systems," *Soft Comput.*, vol. 19, no. 11, pp. 3071–3081, 2015.

[32]S. Q. Salih, A. A. Alsewari, B. Al-Khateeb, and M. F. Zolkipli, "Novel Multi-swarm Approach for Balancing Exploration and Exploitation in Particle Swarm Optimization," in *Recent Trends in Data Science and Soft Computing*, 2019, pp. 196–206.

[33]S. Q. Salih and A. A. Alsewari, "Solving large-scale problems using multi-swarm particle swarm approach," *Int. J. Eng. Technol.*, vol. 7, no. 3, pp. 1725–1729, 2018.

[34]R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on*, 1995, pp. 39–43.

[35]S. Kumar, X. Gao, I. Welch, and M. Mansoori, "A Machine Learning based Web Spam Filtering Approach," in *Advanced Information Networking and Applications (AINA), 2016 IEEE 30th International Conference on*, 2016, pp. 973–980.

[36]UCI Machine Learning Repository, "Spambase Dataset," *University of California, School of Information and Computer Science*. .

[37]K. R. D and P. Visalakshi, "A Hybrid ACO Based Feature Selection Method for Email Spam Classification," *WSEAS TRANSACTIONS on COMPUTERS*, vol. 14. pp. 171–177, 2015.

[38]Geetika, H. Aggarwal, and G. Singh, "A Hybrid Artificial Bee Colony and Support Vector Machine Classifier for Spam Email Detection," *Int. J. Control Theory Appl.*, vol. 10, no. 40, 2016.

[39]D. Karthika Renuka, P. Visalakshi, and T. Sankar, "Improving E-Mail Spam Classification using Ant Colony Optimization Algorithm," *Int. J. Comput. Appl.*