IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 12, 2020, accepted June 4, 2020, date of publication June 19, 2020, date of current version July 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3003569

Particle Swarm Optimization-Based Feature Weighting for Improving Intelligent Phishing Website Detection

WALEED ALI AND SHARAF MALEBARY, (Member, IEEE)
Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
Corresponding author: Waleed Ali (waleedalodini@gmail.com)

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant No. (DF-438-830-1441).

ABSTRACT Over the last few years, web phishing attacks have been constantly evolving causing customers to lose trust in e-commerce and online services. Various tools and systems based on a blacklist of phishing websites are applied to detect the phishing websites. Unfortunately, the fast evolution of technology has led to the born of more sophisticated methods when building websites to attract users. Thus, the latest and newly deployed phishing websites; for example, zero-day phishing websites, cannot be detected by using these blacklist-based approaches. Several recent research studies have been adopting machine learning techniques to identify phishing websites and utilizing them as an early alarm method to identify such threats. However, the important website features have been selected based on human experience or frequency analysis of website features in most of these approaches. In this paper, intelligent phishing website detection using particle swarm optimization-based feature weighting is proposed to enhance the detection of phishing websites. The proposed approach suggests utilizing particle swarm optimization (PSO) to weight various website features effectively to achieve higher accuracy when detecting phishing websites. In particular, the proposed PSO-based website feature weighting is used to differentiate between the various features in websites, based on how important they contribute towards recognizing the phishing from legitimate websites. The experimental results indicated that the proposed PSO-based feature weighting achieved outstanding improvements in terms of classification accuracy, true positive and negative rates, and false positive and negative rates of the machine learning models using only fewer websites features utilized in the detection of phishing websites.

INDEX TERMS Feature weighting, machine learning, particle swarm optimization, phishing website.

I. INTRODUCTION

In recent years, the number of web users who use online services, online shopping, and e-banking has been increasing rapidly due to flexibility, comfort, and ease of use. The huge growth of using online services and e-business has motivated numerous phishers and cyber attackers in developing and publishing deceptive and phishing websites [1]–[3] to gain confidential and financial information of web users. Thus, web phishing attacks have become a serious problem for both web users and commercial websites.

The phishers can steal users' information using many innovative methods such as IMs (Instant Messaging), forums, black hat SEO (Search Engine Optimization), key-loggers, Internet relay chat, trojans, and screen captures [4]–[8]. The DNS-based phishing or pharming is another type of phishing attack, in which hackers alter the host's files or domain name system. Consequently, the requests for URLs return a false address and subsequent communications are directed to a phishing website [4]–[6]. The spear-phishing based on email is another serious version of the targeted phishing attack, in which phishers send fake emails, impersonating business officials, to specific users within an organization to harvest crucial business-related details [9].

In this paper, we focus on one of the most dangerous phishing attacks, known as phishing websites-based attack. In this attack, the phisher develops a fake or phishing website that looks like a replica of a legitimate website to convince

The associate editor coordinating the review of this manuscript and approving it for publication was Nilanjan Dey.

the victims to give out their confidential and financial information. Then, the phisher tries to attract many victims to this phishing website either through email, social networks, or advertisements on other websites. When the victims land on the phishing website, they get deceived that it is a legitimate website and give out their confidential and financial information [3], [7], [10].

In the past few years, numerous web phishing attacks have been continuously developing causing less trust of customers in online commerce and business. FIGURE 1 shows the number of unique phishing web sites detected by APWG (Anti-Phishing Working Group) for three quarters of 2019 as reported in [11].

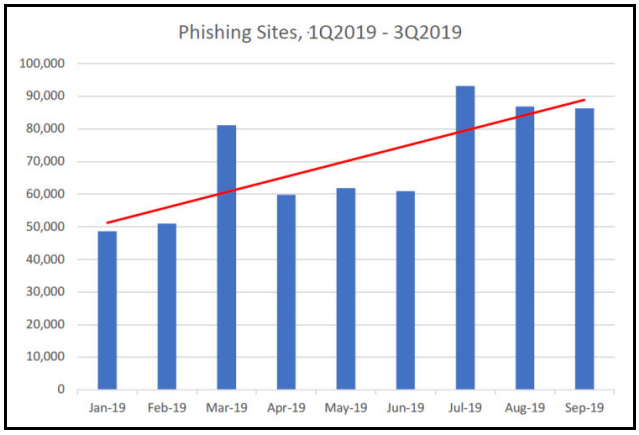

FIGURE 1. The number of unique phishing web sites detected by APWG for three quarters of 2019 [11].

Many blacklist-based tools and strategies are developed to identify web phishing attacks and alert users when they land on a phishing website [9]. The browser toolbars and DNS blacklist are two popular blacklist-based tools and strategies employed to identify phishing websites. The browser toolbar performs filtering URLs from the address bar and then issues a caution if this URL is available in a blacklist. Numerous safe browsing toolbars have been developed to work with common browsers [12] such as Chrome [13], Firefox, Safari, and Internet Explorer. On the other hand, the DNS blacklist approach is based on lists of known phishing sites that providers have. These lists get updated regularly with new phishing websites, and support query methods for users; for example, SORBS [14], URIBL [15], and SURBL [16].

The blacklist-based tools and approaches can effectively detect the phishing websites if they are available in the phishing websites database. However, the latest phishing websites deployed newly, especially zero-day phishing websites, cannot be detected by using these blacklist-based tools and approaches [3], [7], [10], [17]–[19].

In order to detect successfully new phishing websites, several recent research studies [3], [7], [10], [17]–[22] have suggested training of some common machine learning techniques based on datasets contain legitimate and phishing websites. After the training phase is successfully completed, the trained classifiers are used in differentiating new phishing websites from legitimate ones.

In order to enhance the detection accuracy and speed up the classifiers used to detect phishing websites, the most significant website features were extracted and selected using methods based on human experience [9], [12], [18], [23] and frequency analysis [7], [17], [24]–[26]. These methods required more effort and considerable time since the feature's assessment was conducted manually or based on frequency analysis of the various features obtained from many websites experimentally or collected from the previous research studies. Alternatively, the best minimal set of phishing website features has been selected automatically using filter-based feature selection methods [7], [8], [27], [28]. The filter methods evaluate the features independently of a specific machine learning algorithm. Thus, some machine learning algorithms with considering filter-based feature selection methods achieved better detection accuracy, while others produced similar or slightly worse detection accuracy of phishing websites. In recent years, feature selection methods based on wrapper approach [10] and genetic algorithm (GA) [29] have been used to obtain the most influential website features to achieve higher detection accuracy of the phishing websites. Although the wrapper and GA based feature selection methods achieved better detection accuracy and outperformed other feature selection methods, they were time-consuming for some machine learning algorithms.

Alternatively, this paper proposes implementing particle swarm optimization (PSO) to produce and assign a weight to each website feature in order to help in increasing the accuracy of phishing website detection with feasible computation and resources. This weight represents the importance and relevance of the feature for phishing website detection. The proposed PSO-based feature weighting method has the following attractive benefits in contrast to other existing feature selection methods:

- The proposed method is based on the great performance of PSO, which is one the most well-known evolutionary algorithms that are not only search for the best solution, but they are also able to evolve solutions to produce the optimal solution.
- Unlike the former and known feature selection methods discussed in the literature, the proposed method employs PSO to weight the website features effectively in order to increase the performance of machine learning. In other words, the proposed PSO-based website feature weighting is used to differentiate between the various features of websites, based on how important they contribute towards recognizing the phishing from legitimate websites.
- The proposed PSO-based feature weighting utilizes PSO, which is a simpler and faster evolutionary algorithm and has fewer parameters compared to GA.
- The proposed PSO-based feature weighting can produce exciting enhancements in the performances of machine

learning classifiers since it is classified under the wrapper approach, which often produces better results than filter techniques.
- Using the proposed PSO-based feature weighting, between 7% and 57% of irrelevant features are removed, and only the remaining features are utilized with the classifiers to detect the phishing websites.
- The machine learning models enhanced by the proposed PSO-based feature weighting can achieve better detection accuracy and outperform the stand-alone machine learning models, and these machine learning models with applying other feature selection methods.
- The machine learning models improved by the proposed PSO-based feature weighting can achieve better-balanced performance in the detection of both phishing and legitimate websites since they perform good detection results in terms of true positive rate, true negative rate, false positive rate, and false negative rate.

The remaining sections of this article are organized as follows. Section II presents and discusses some of the recent existing works that applied feature selection methods with machine learning techniques to enhance the detection of phishing websites. Section III reviews the features categories of phishing websites and the relevant features of each category. The basic concepts of feature weighting and particle swarm optimization are described in Sections IV and V, respectively. In Section VI, the PSO-based feature weighting approach suggested for improving the detection of phishing websites is presented and explained. Section VII discusses and deliberates the detection performance of popular machine learning models with and without the proposed PSO-based feature weighting. Besides, Section VII compares the performance of the proposed PSO-based feature weighting with common feature selection methods. Eventually, Section VIII concludes and suggests future work of this study.

II. RELATED WORK

In this section, we review some of the recent existing works that applied some feature selection methods with machine learning techniques to enhance the detection of phishing websites. Generally, the feature selection methods utilized in detecting phishing websites can be categorized into four categories: frequency analysis-based feature selection, filter-based feature selection, wrapper-based feature selection, and evolutionary algorithm-based feature selection.

Many research works have utilized frequency analysis-based feature selection to find significant features to improve the performance of intelligent methods in recognizing the legitimate from phishing websites. In [26], the authors assessed many websites' features using a software tool to compute each feature frequency, which represents the feature importance. In [17], seventeen significant features were identified based on frequency analysis. The selected features were used to train self-structuring neural networks in order to distinguish between phishing websites and legitimate ones. In a similar way to [17], [24] analyzed the frequency of websites' features to select the most popular features of websites. Then, rule-based data mining classification models were trained based on the selected website features to recognize the new phishing websites. Find function was exploited by [25] to investigate the most substantial features that exist frequently in numerous websites. Neuro-Fuzzy was then trained with the best five features to detect the phishing websites through an online transaction.

Alternatively, several recent existing works demonstrated that the filter-based feature selection techniques enhanced noticeably the performance of intelligent phishing detection approaches. In [7], the authors exploited both frequency analysis and Chi-Square to select a minimal set of relevant websites features from the original features. Based on the selected web site's features, a MCAC (Multi-label Classifier based Associative Classification) model was trained and developed to distinguish the phishing websites from legitimate ones. Information Gain (IG), Chi-square, and Correlation Feature Set were employed by [30] to find the most significant website's features in order to enhance the detection accuracy of phishing websites for some rule-based classification machine learning algorithms: C4.5, RIPPER, and PART. In [8], the authors suggested using the IG, Chi-square, and Correlation Features Set (CFS) to reduce the data dimensionality and select the minimal set of important features. Then, four rule-based classification algorithms (OneRule, JRip, Part, and J48) were trained after applying feature selection methods in order to maximize the detection rate of phishing emails.

The results in the studies mentioned earlier showed that some machine learning algorithms based on filter-based feature selection achieved better detection accuracy of phishing websites and emails. However, other machine learning algorithms that applied the filter methods may suffer from relatively poor performance since the filter-based feature selection methods utilize statistical measures to rate each feature independently of a specific machine learning algorithm.

The wrapper feature selection method coupled with the best-first forward searching method was also applied in phishing email classification by [31] and then compared against IG, Relief-F, and CFS. In [31], the authors demonstrated that the wrapper feature selection method outperformed IG, Relief-F, and CFS. To identify phishing websites accurately, the most significant websites' features were selected in [10] by using the wrapper-based feature selection. Accordingly, the training dataset with the selected features was used to train RBFN, SVM, NB, C4.5, kNN, and RF. Results indicated that RBFN, SVM, NB, C4.5, kNN, and RF with the considering wrapper-based feature selection accomplished better detection accuracy compared to theses machine learning classifiers based on IG and PCA (Principal Component Analysis). However, the wrapper-based feature selection depends on the machine learning algorithm itself and may be computationally expensive.

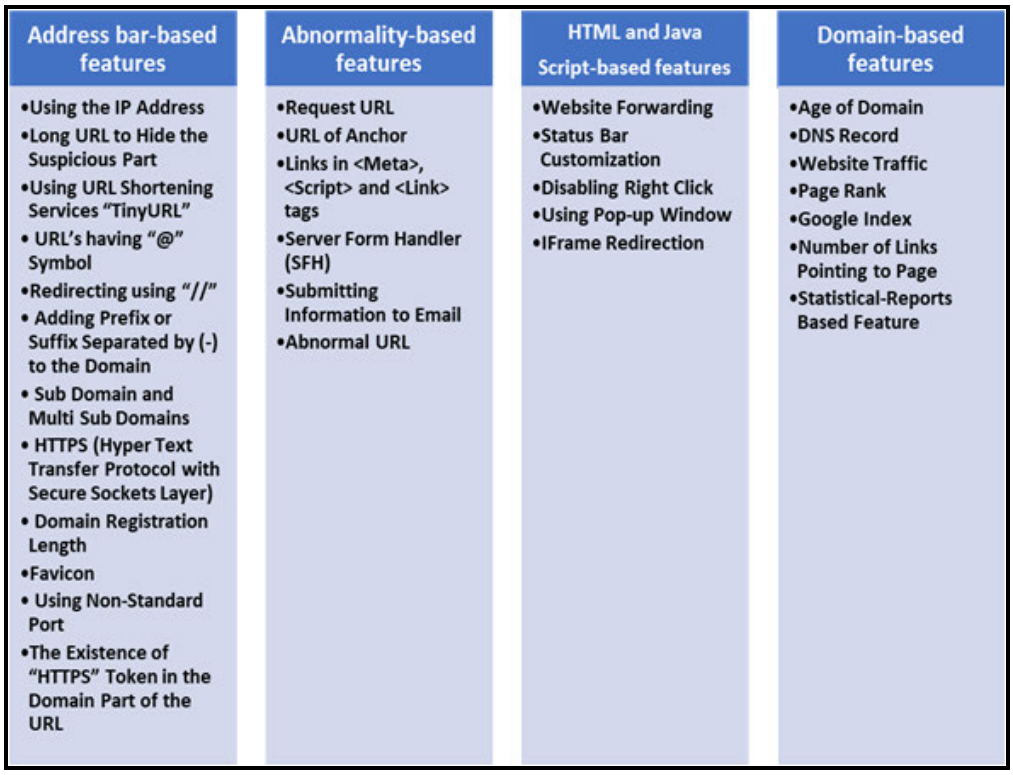

FIGURE 2. The relevant features of phishing websites.

Recently, genetic algorithm-based feature selection was used in [29] to find more relevant features in order to enhance the detection accuracy of the machine learning model in phishing websites detection. Although the machine learning techniques with applying GA-based feature selection performed better detection accuracy compared to the same machine learning techniques with other feature selection methods, GA-based feature selection required a longer time for some machine learning algorithms.

III. OVERVIEW OF RELEVANT FEATURES OF PHISHING WEBSITES

This section reviews the most popular websites' features that are employed in the field of intelligent phishing website detection to differentiate between phishing and legitimate websites.

Unlike the blacklist-based conventional approach, the success of an intelligent detection approach of the phishing website is extremely based on extracting common features from websites to train machine learning models effectively in order to recognize phishing websites [3], [7], [10], [17]–[19].

Some websites' features are more significant and relevant than others in contributing to recognizing the phishing from legitimate websites. Thus, extracting these discriminative websites' features plays an extremely important role in increasing the detection accuracy of the phishing websites. Numerous popular websites' features have been investigated and assessed in the literature to identify the most significant features in phishing websites detection. In [26], [32], the authors adopted four categories of relevant features of phishing websites: HTML and JavaScript-based features, address bar-based features, domain-based features, and abnormality-based features. FIGURE 2 shows the features categories of phishing websites and the relevant features of each category.

IV. FEATURE WEIGHTING

In recent years, a huge amount of high-dimensional data in numerous fields has led to a computational challenge, which makes machine learning algorithms inapplicable or difficult to be applied in many real-world problems [33]–[35]. Feature selection and weighting are two common data reduction techniques, which can be used to face the problem of the high-dimensional data.

The feature selection aims to eliminate the unnecessary and irrelevant features and choose only the most relevant features to improve classification performance and accuracy of machine learning models.

There are two popular approaches used in feature selection namely filter and wrapper. The filter approach uses statistical measures or criteria to evaluate the relevance of the features based on the essential characteristics of the training data without involving any machine learning algorithm. On the other hand, the wrapper approach directly uses a machine-learning algorithm to evaluate the goodness of features. Thus, wrapper methods generally perform a higher classification accuracy

than filter methods, but they are computationally more intensive than filter methods.

Unlike feature selection, feature weighting is a popular strategy used to find the optimal weight for each feature, which represents the importance of the feature for the classification decision. Thus, the feature weighting is utilized to assign higher and lower weights to differentiate between the most important and the less important features, respectively. Furthermore, the feature weighting can also reduce the number of features by giving zero weights to irrelevant and redundant features. Accordingly, a machine learning model is built and trained based on the training data with a smaller set of weighted features.

Like feature selection, there are filter and wrapper approaches in the feature weighting. The filter feature weighting methods compute the weights of the features based on the general characteristics of data independently of a specific machine learning algorithm. On the other hand, the wrapper feature weighting methods optimize the weights of the features based on a particular machine learning classifier used as fitness objective function during the process of feature evaluation [36].

Many recent research studies demonstrated that the feature weighting usually performs better classification accuracy compared to the feature selection when both are used to tackle the same problem [33], [35], [37]–[39]. This is because the feature selection can be regarded as a particular case of feature weighting since the feature selection can assign only two weights (binary value) to features – either 1 if the feature is selected or 0 if it is not selected. More generally, the feature weighting can assign weights with a real number, usually in the interval [0,1], to differentiate between the most important and the less important features.

V. PARTICLE SWARM OPTIMIZATION

Particle swarm optimization (PSO) was firstly invented by Kennedy and Eberhart [40] as a population-based evolutionary algorithm to simulate the cooperative behavior of birds flocking in finding the food. The main advantage of PSO over other optimization algorithms is its ability to achieve fast convergence in numerous complicated optimization problems. Furthermore, PSO has several attractive advantages such as simplicity with fewer mathematical equations and having fewer parameters in implementation [41]–[45].

In PSO, every candidate solution is represented by a particle, which flies with a specific velocity in a swarm or population representing a group of possible solutions.

Initially, all particles in the swarm are generated by assigning random positions and velocities. Then, each particle in the swarm adjusts its position and velocity dynamically according to its own flying experience and its companions' flying experience.

In each PSO iteration, each particle keeps a record of its previous best position (pbest) and can access to the recorded global best position (gbest). Accordingly, each particle adjusts its position and velocity based on pbest and gbest using Equations (1) and (2) in order to obtain the optimal solution in the swarm.

$$x_{id}^{t+1} = x_{id}^{t} + v_{id}^{t+1} \tag{1}$$

$$v_{id}^{t+1} = w * v_{id}^{t} + c_1 * r_1 * (p_{id} - x_{id}^{t}) + c_2 * r_2 (p_{gd} - x_{id}^{t}) \tag{2}$$

where x_{id}^{t} and v_{id}^{t} represent the position and velocity of particle i, respectively, at the tth iteration while $d = 1, 2, 3 \ldots D$ (D is the dimensionality of search space). The pbest and gbest are denoted by p_{id} and p_{gd}, respectively. c_1 and c_2 are positive constants and denoted for learning rates, which are frequently set to 2.0. They represent the weighting of the stochastic acceleration terms that pull each particle towards its pbest and gbest positions. w represents the inertia weight whereas r_1 and r_2 are randomly set to real numbers in the interval [0, 1].

VI. METHODOLOGY

In this section, the proposed PSO-based feature weighting approach suggested to improve the phishing website prediction is presented and explained. FIGURE 3 illustrates a methodology of the proposed intelligent phishing website detection based on PSO-based feature weighting. The methodology consists of two main phases: training phase and detection phase. Each phase will be explained in detail in the following.

In the training phase, a set of phishing and legitimate websites is used to train and build an intelligent detection model. Initially, 12 features of address bar-based category, 6 features of abnormality-based category, 5 features of HTML and JavaScript-based category, and 7 features of domain-based category are extracted from 11055 websites. Since these extracted features have different importance and can contribute differently to phishing websites detection, the extracted website features are weighted using the proposed PSO-based feature weighting. Accordingly, the machine learning algorithms are trained using the websites' features weighted by PSO to precisely detect the phishing websites.

Unlike the feature selection which entirely excludes the less important features, the feature weighting aims to assign lower weights to less influential features, and higher weights to more significant features in order to improve the effectiveness of machine learning models.

In the proposed PSO-based feature weighting, the best weights of website features are heuristically generated using PSO to maximize the performance of phishing website detection. The PSO-based feature weighting adopted to contribute toward enhancing the phishing website detection is shown in FIGURE 4.

To apply the proposed PSO-based feature weighting, it is required to encode the weights of the features in particles which are real values in the interval [0,1], representing the possible weights of website features. The dimensionality of each particle is the number of features to be weighted

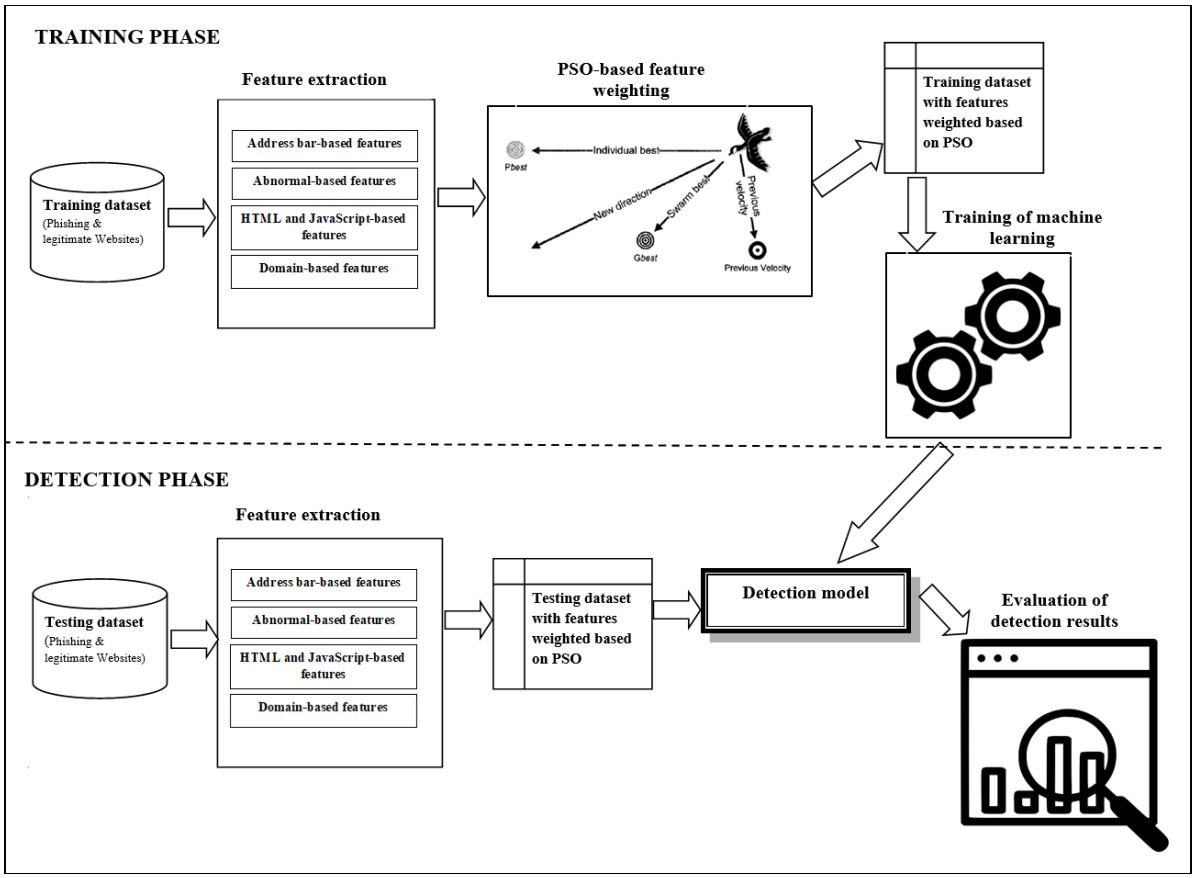

FIGURE 3. The proposed intelligent phishing website detection based on PSO-based feature weighting.

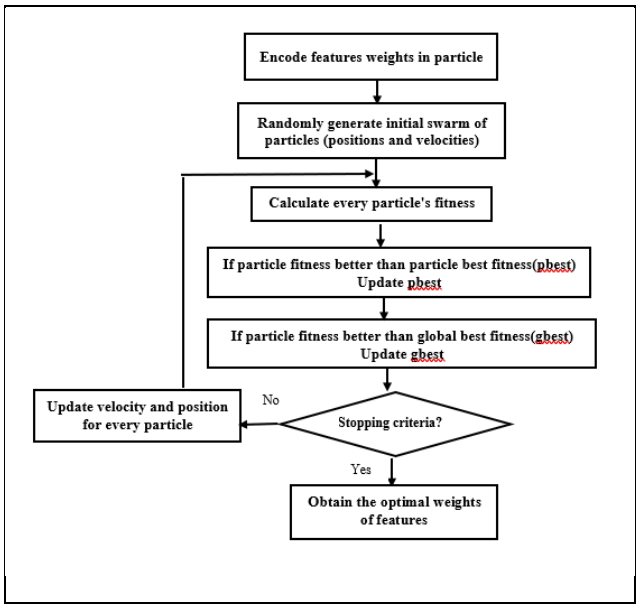

FIGURE 4. The PSO-based feature weighting adopted to improve the phishing websites detection.

using PSO. An example of encoding weights for a set of n features in PSO particle in the proposed PSO-based feature weighting is shown in FIGURE 5.

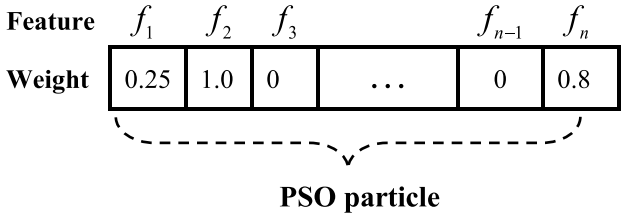

FIGURE 5. An example of encoding weights for a set of n features in PSO particle.

The example in FIGURE 5 illustrates the second and nth features are the most significant among other features since they have the highest weights. On the other hand, the first feature is less relevant than the second and nth features, while the third and n-1th features are considered redundant and irrelevant features and can be removed during the training of intelligent detection models.

In the proposed PSO-based feature weighting, a swarm of particles that represent the possible sets of features weights is initialized randomly. Each particle has a random position and velocity. The position of each particle in the swarm, which represents a set of the features' weights, is encoded with real values between zero and one. After generating an initial swarm of particles, each particle's fitness is then computed

and evaluated. In the proposed PSO-based feature weighting, the classification accuracy is the particle's fitness. In other words, the PSO assesses the fitness of each particle by training the machine learning using the training dataset with the features weighted with that particle and then computing the classification accuracy to be used as that particle's fitness. The objective of PSO fitness evaluation is to find the personal best position(pbest) for each particle and the global best position (gbest) for the whole swarm. Then, the current pbest and gbest will become the pbest and gbest if their fitness values (correct classification rates) are higher than fitness values of previous pbest and gbest, respectively. Accordingly, every particle changes its velocity and position based on the new pbest and gbest using Equations (1) and (2). This process is repeated until PSO finishes the maximum number of iterations. Eventually, PSO returns the gbest, which represents a set of the ideal features' weights available in the swarm.

After feature weighting based on PSO, six popular machine learning algorithms, which were commonly used in literature, are trained using the training dataset with the features weighted by PSO. In this paper, back-propagation neural network (BPNN), support vector machine (SVM), k-Nearest neighbor (kNN), decision tree (C4.5), random forest (RF), and naïve Bayes classifier (NB) are trained using the training dataset of features weighted by PSO. The trained models are subsequently created and saved to be used in the detection phase in order to detect the new phishing websites.

In the detection phase, new websites are collected and used as a testing dataset to assess the performance of the phishing website detection models created in the training phase. In a similar manner to extract the features of the websites in the training phase, it is required to extract the essential features of the testing dataset: HTML and JavaScript-based features, address bar-based features, domain-based features, and abnormality-based features. The optimal weights obtained by PSO in the training phase are then utilized to weight features of the testing dataset to enhance the detection accuracy of phishing websites. Accordingly, the features weighted by PSO are used as input features of the phishing website detection models, which created in the training phase, to classify whether the website is phishing or not. Eventually, the performances of phishing website detection models based on the suggested feature weighting using PSO are evaluated and compared to the stand-alone phishing website detection models to analyze the enhancement obtained in detecting the new phishing websites.

VII. EXPERIMENTAL RESULTS AND DISCUSSION

A. DATASET COLLECTION

In this paper, we conducted the experiments with the phishing websites dataset available for free use in UCI Machine Learning Repository [46] in order to evaluate the performance of the proposed PSO-based feature weighting approach suggested to improve the phishing website detection. In this phishing websites dataset, there are 4898 phishing websites and 6157 legitimate websites out of 11055 websites. The essential characteristics of the phishing websites datasets used in the experiments and evaluation are summarized in TABLE 1.

TABLE 1. The essential characteristics of phishing websites dataset used in the experiments.

Characteristic	Description
Websites features	12 features of address bar-based category, 6 features of abnormality-based category, 5 features of HTML and JavaScript-based category, and 7 features of domain-based category. Refer to FIGURE 2 to know features names
# Features	30
Classes	Legitimate or phishing website
# Classes	2
# Websites	11055
#Phishing websites	4898
Percentage of phishing websites	44 %
#Legitimate websites	6157
Percentage of legitimate websites	56 %

B. EVALUATION MEASURES

In our experiments, the proposed method was implemented and evaluated using RapidMiner, which is a popular data science and machine learning platform. We used five commonly used measures to evaluate the performance of the PSO-based feature weighting suggested to enhance phishing website detection based on machine learning techniques. TABLE 2 shows the confusion matrix for the problem of phishing website detection, while TABLE 3 describes the five popular classification metrics used in this paper to evaluate the proposed PSO-based feature weighting.

TABLE 2. Confusion matrix for the problem of phishing website detection.

	Classified as phishing	Classified as legitimate
Phishing website	True Positive (TP)	False Negative (FN)
Legitimate website	False Positive (FP)	True Negative (TN)

The five popular metrics used are classification accuracy, true positive rate (TPR) or sensitivity, true negative rate (TNR) or specificity, false positive rate (FPR), and false negative rate (FNR). The classification accuracy is the most popular classification measure, which is used to show the percentage of websites that are correctly classified. The TPR is the number of phishing websites correctly classified as

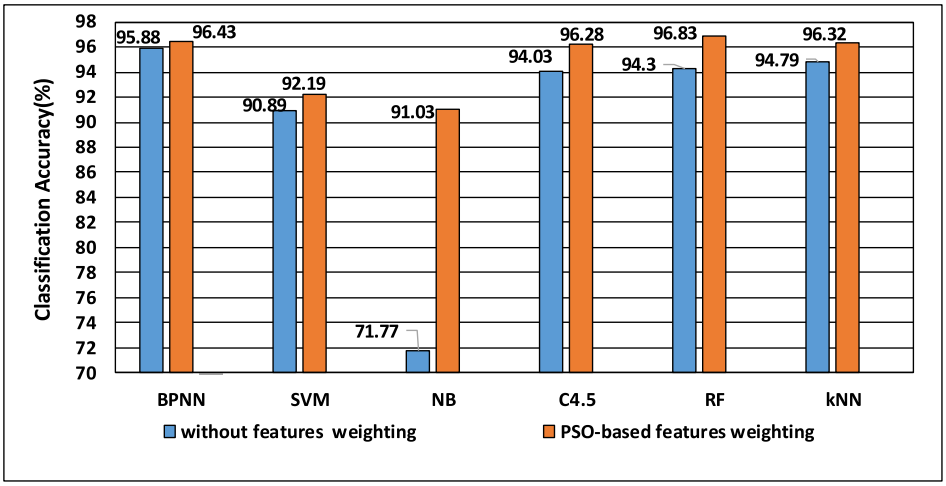

FIGURE 6. The average values of classification accuracy achieved by machine learning classifiers before and after implementing the proposed PSO-based feature weighting.

TABLE 3. The most popular classification measures used to evaluate the proposed PSO-based feature weighting.

Measure name	Formula (%)
Classification accuracy (CCR)	$CCR = \frac{TP + TN}{TP + FP + FN + TN} \times 100$
True positive rate (TPR)	$TPR = \frac{TP}{TP + FN} \times 100$
True negative rate (TNR)	$TNR = \frac{TN}{TN + FP} \times 100$
False positive rate (FPR)	$FPR = \frac{FP}{FP + TN} \times 100$
False negative rate (FNR)	$FNR = \frac{FN}{FN + TP} \times 100$

phishing divided by the total phishing websites. The TNR is the number of legitimate websites correctly classified as legitimate out of total legitimate websites. The FPR is the number of legitimate websites misclassified as phishing divided by the total legitimate websites. The FNR is the number of phishing websites misclassified as legitimate divided by the total phishing websites. For better performance, the proposed phishing website detection should achieve high classification accuracy, TPR, and TNR, and produce low FPR and FNR.

In order to precisely assess the proposed method, common machine learning techniques with and without PSO-based feature weighting were validated using 10-fold cross-validation. The performances of Back-propagation neural network (BPNN), support vector machine (SVM), k-Nearest neighbor (kNN), decision tree (C4.5), random forest (RF), and the naïve Bayes classifier (NB) were evaluated before and after implementing the proposed feature weighting based on PSO. Furthermore, we compared the performance of the proposed PSO-based feature weighting with other common feature selection methods used in literature for detecting phishing websites.

C. COMPARISON OF MACHINE LEARNING CLASSIFIERS BEFORE AND AFTER APPLYING THE PROPOSED PSO-BASED FEATURE WEIGHTING

In all experiments, we select the best parameters of PSO used in the proposed PSO-based feature weighting by a trial-and-error basis in order to produce the best detection results. The best parameters of PSO used in the proposed PSO-based feature weighting are shown in TABLE 4.

TABLE 4. Parameters of PSO used with the PSO-based feature weighting.

Parameter	Value
Number of particles	20
Maximum iterations (generations)	50
Local best weight(c1)	2
Global best weight(c2)	2
Inertia weight(w)	1
Stop condition	Maximum number of iterations

In this section, we discuss the performance enhancement of machine learning techniques in detecting the phishing websites obtained after applying the proposed PSO-based feature weighting. FIGURE 6 shows the performance in terms of classification accuracy of BPNN, SVM, NB, C4.5, RF, and kNN used to detect the phishing websites before and

TABLE 5. A comparison of average values of TPR, TNR, FPR, and FNR achieved by machine learning classifiers before and after applying the proposed PSO-based feature weighting.

	Measures	Without feature weighting	PSO-based feature weighting
BPNN	TPR	95.06	95.57
	TNR	96.54	97.11
	FPR	3.46	2.89
	FNR	4.94	4.43
SVM	TPR	86.38	89.1
	TNR	94.48	94.66
	FPR	5.52	5.34
	FNR	13.62	10.9
NB	TPR	99.51	86.81
	TNR	49.7	94.38
	FPR	50.3	5.62
	FNR	0.49	13.19
C4.5	TPR	93.9	95.3
	TNR	94.14	97.06
	FPR	5.86	2.94
	FNR	6.1	4.7
RF	TPR	90.83	95.37
	TNR	97.06	98
	FPR	2.94	1.2
	FNR	9.17	4.63
kNN	TPR	93.75	96.14
	TNR	95.61	96.46
	FPR	4.39	3.54
	FNR	6.25	3.86

after applying the proposed PSO-based feature weighting. In addition, the performances in terms of TPR, TNR, FPR, and FNR of BPNN, SVM, NB, C4.5, RF and kNN before and after applying the proposed PSO-based feature weighting are presented in TABLE 5.

In terms of classification accuracy, FIGURE 6 shows that the highest classification accuracy was achieved by BPNN (95.88%), kNN (94.79%), RF (94.3%), and then C4.5(94.033%). SVM performed good classification accuracy (90.89%) but less than BPNN, kNN, RF, and C4.5, while NB produced the worst classification accuracy (71.77%) in the phishing website detection. Furthermore, FIGURE 6 demonstrated that the performances of BPNN, SVM, NB, C4.5, RF, and kNN were outstandingly enhanced after applying the proposed PSO-based feature weighting. FIGURE 6 shows that the proposed PSO-based feature weighting enhanced the classification accuracies of BPNN, SVM, NB, C4.5, RF and kNN from 95.88%, 90.89%, 71.77%, 94.03%, 94.3% and 94.79% to 96.43%, 92.19%, 91.03%, 96.28%, 96.83% and 96.32%, respectively. This was primarily due to the capability of the proposed PSO-based feature weighting to successfully weight the website features used for enhancing phishing website detection.

In addition to the classification accuracy, TABLE 5 shows TPR, TNR, FPR, and FNR of machine learning classifiers before and after applying the proposed PSO-based feature weighting.

As seen in TABLE 5, the TPRs of classifiers after applying the proposed PSO-based feature weighting were remarkably enhanced, except for NB,compared to the stand-alone classifiers used for detecting the phishing websites.

In terms of TNR, TNRs achieved by all machine learning classifiers after applying the proposed PSO-based feature weighting were higher than TNRs of stand-alone classifiers. The higher performance in both TPR and TNR indicated that both phishing and legitimate websites were accurately detected by the machine learning models improved by using the proposed PSO-based feature weighting.

In terms of FPR and FNR, FPR and FNR are equivalent of 1 - TNR and 1 – TPR, respectively. A better phishing detection model should produce lower FPR and FNR [47], [48]. Furthermore, there is a trade-off between TPR and FPR [47], [48]. As a consequence of higher TPR and TNR, most of the machine learning models improved by the proposed PSO-based feature weighting achieved better FPR and FNR compared to the stand-alone machine learning models.

It can be observed from TABLE 5, unbalanced performance in the detection of both phishing and legitimate websites was produced by NB since NB achieved the best TPR and FNR and the worst TNR and FPR. That means NB classified inaccurately most of websites as phishing website. Although the stand-alone NB classifier accomplished the highest TPR (99.51%), it produced the worst FPR (50.3%). That means 50 % of legitimate websites were classified as phishing websites. On the other hand, the NB classifier and other machine learning models improved by the proposed PSO-based feature weighting achieved better-balanced performance in the detection of both phishing and legitimate websites since they performed good detection results in terms of TPR, FNR, TNR, and FPR.

In order to validate the quality of the proposed method, we plot the receiver operating characteristics (ROC) graph for the machine learning classifiers improved by the proposed PSO-based feature weighting. The ROC graph illustrates the relative tradeoffs between benefits (true positives) and costs (false positives).

The ROC curve is drawn by plotting TPR against FPR as y and x axes, respectively, at various threshold settings. The perfect classification model would produce a point in the upper left corner or coordinate (0,1) of ROC, representing 100% TPR and zero FPR (no false positives). The performance of classification model will be better if the ROC curve drawn by this classifier is above the random classifier, which represents the curve formed by the (0, 0) and (1, 1) connections. FIGUREs 7-12 show the ROC curves of BPNN, SVM, NB, C4.5, RF, and kNN improved by the

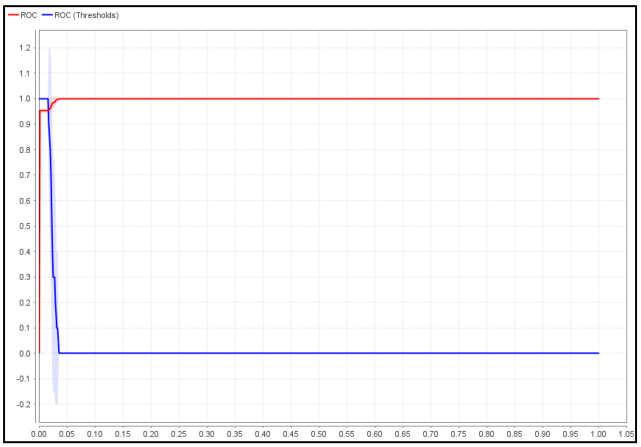

FIGURE 7. ROC curve of kNN improved by the proposed PSO-based feature weighting.

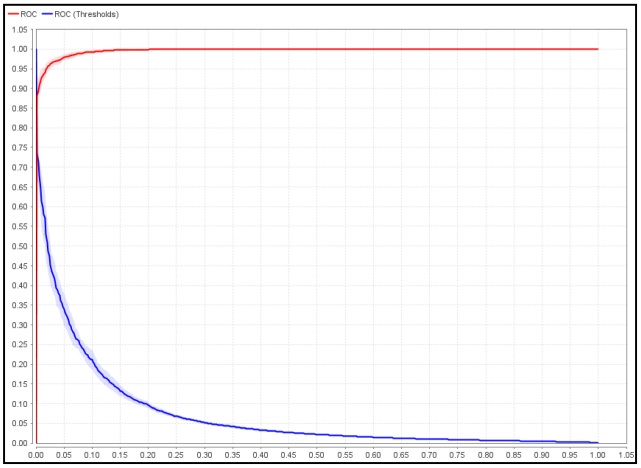

FIGURE 8. ROC curve of RF improved by the proposed PSO-based feature weighting.

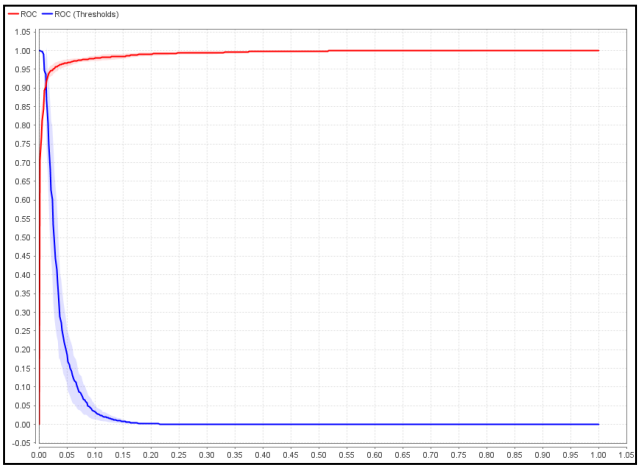

FIGURE 9. ROC curve of BPNN improved by the proposed PSO-based feature weighting.

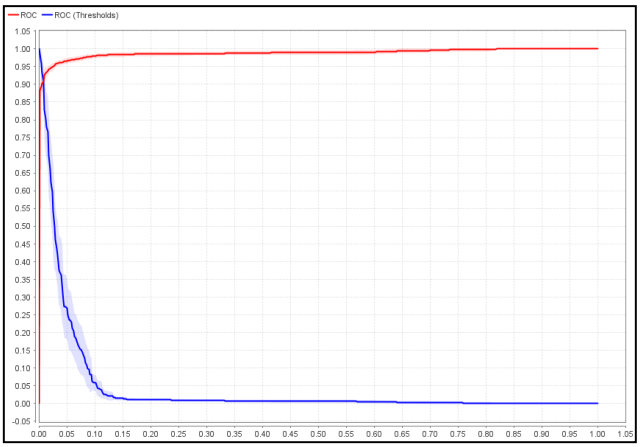

FIGURE 10. ROC curve of C4.5 improved by the proposed PSO-based feature weighting.

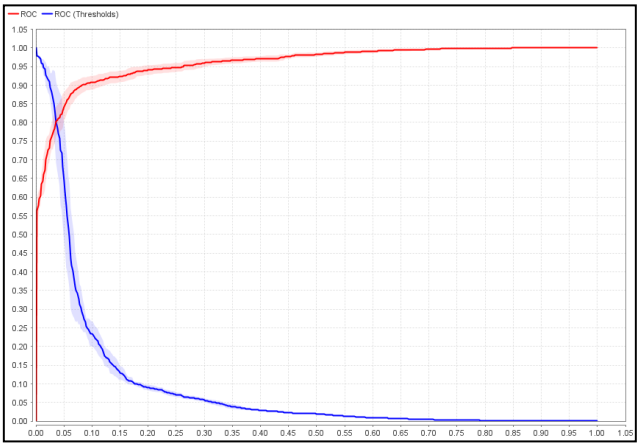

FIGURE 11. ROC curve of NB improved by the proposed PSO-based feature weighting.

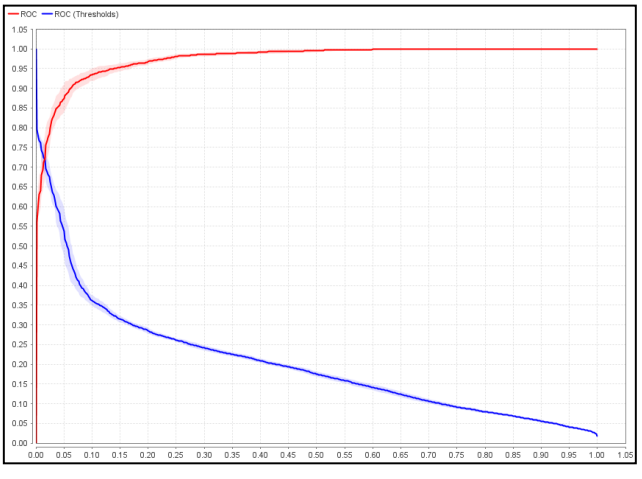

FIGURE 12. ROC curve of SVM improved by the proposed PSO-based feature weighting.

proposed PSO-based feature weighting. As can be observed from FIGUREs 7-12, BPNN, SVM, NB, C4.5, RF, and kNN improved by the proposed PSO-based feature weighting have good classification performance since the curves drawn by these classifiers improved by the proposed PSO-based feature weighting are above the random classifier curve.

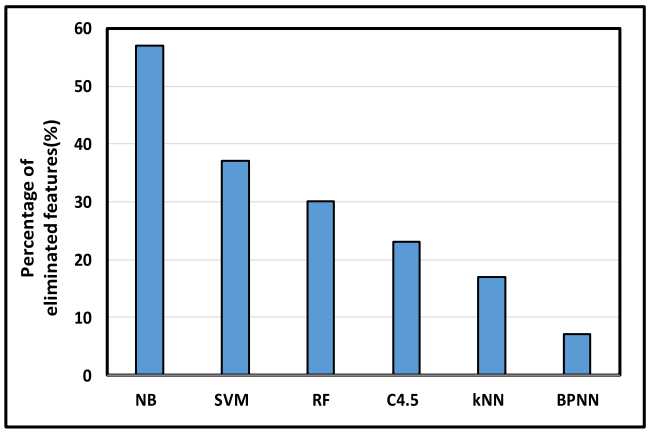

FIGURE 13. The percentage of irrelevant websites features that were removed by the proposed PSO-based feature weighting.

Furthermore, we calculated the area under ROC curve (AUC), which is a common classification measure equivalent to the probability that the classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. All classifiers improved by the proposed PSO-based feature weighting archived high AUC values between 0.96 and 0.999. In particular, the AUC values of kNN, RF, BPNN, C4.5, SVM, and NB that applied the proposed PSO-based feature weighting are 0.999, 0.996, 0.991, 0.989, 0.975, and 0.960, respectively.

D. REDUCTION OF IRRELEVANT FEATURES USING THE PROPOSED PSO-BASED FEATURE WEIGHTING

In addition to enhancing the performances of the classifiers, the proposed PSO-based feature weighting was able to produce the optimal weight to each website feature based on its importance and influence on the detection decision of phishing websites. The significant features had higher weights, while the less relevant features were weighted with lower weights by PSO. Moreover, the proposed PSO-based feature weighting contributed to eliminating the irrelevant and redundant features and thus decreasing the number of features used in the training of machine learning.

FIGURE 13 shows that the proposed PSO-based feature weighting omitted between 7% and 57% of irrelevant features and only the remaining features were utilized to train machine learning classifiers in order to identify the phishing websites. Thus, the classifiers based on PSO-based feature weighting achieved better detection results using only fewer features since they were able to exclude unnecessary and redundant features.

E. COMPARISON OF THE PROPOSED PSO-BASED FEATURE WEIGHTING AGAINST THE EXISTING FEATURE SELECTION METHODS

In this section, we compare the performances of machine learning models improved by the proposed PSO-based feature weighting against these models that applied other feature selection methods used in the phishing website detection such as Chi-square and Information Gain (IG) [7], [8], [30], Wrapper-based feature selection [10], GA-based features selection [29], and GA-based features weighting [29].

In terms of classification accuracy, FIGURE 14 shows a comparison of the detection accuracy of the machine learning models that applied all methods: the proposed PSO-based feature weighting, IG, Chi-square, Wrapper feature selection, GA-based features selection, and GA-based features weighting. As can be observed from FIGURE 14, the proposed

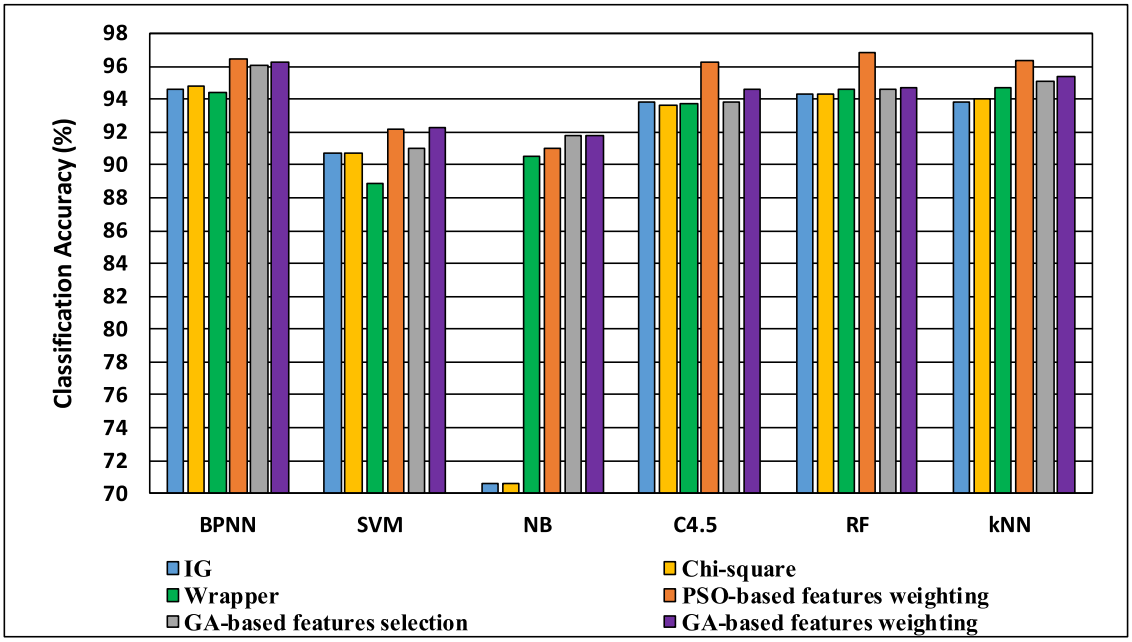

FIGURE 14. A comparison of the average values of classification accuracy achieved by the machine learning classifiers improved by the proposed PSO-based feature weighting against other feature selection and weighting methods.

TABLE 6. A comparison of the average values of TPR, TRN, FPR and FNR achieved by the machine learning classifiers improved by the proposed PSO-based feature weighting against other feature selection and weighting methods.

	Measures	IG	Chi-square	Wrapper	PSO-based feature weighting	GA-based features selection	GA-based features weighting
BPNN	TPR	93.63	93.96	92.34	95.57	93.4	95.17
	TNR	95.45	95.53	96.07	97.11	98.21	97.13
	FPR	4.55	4.47	3.93	2.89	1.79	2.87
	FNR	6.37	6.04	7.66	4.43	6.6	4.83
SVM	TPR	86.08	86.16	85.69	89.1	86.04	86.86
	TNR	94.38	94.28	91.44	94.66	94.91	96.64
	FPR	5.62	5.72	8.56	5.34	5.09	3.36
	FNR	13.92	13.84	14.31	10.9	13.96	13.14
NB	TPR	99.69	99.71	88.55	86.81	89.79	89.79
	TNR	47.43	47.51	92.11	94.38	93.45	93.45
	FPR	52.57	52.49	7.89	5.62	6.55	6.55
	FNR	0.31	0.29	11.45	13.19	10.21	10.21
C4.5	TPR	93.06	93.02	92.38	95.3	91.42	94.15
	TNR	94.48	94.14	94.77	97.06	95.78	94.91
	FPR	5.52	5.86	5.23	2.94	4.22	5.09
	FNR	6.94	6.98	7.62	4.7	8.58	5.85
RF	TPR	90.96	91.02	92.38	95.37	91.49	92.04
	TNR	97.06	96.96	96.43	98	97.02	96.91
	FPR	2.94	3.04	3.75	1.2	2.98	3.09
	FNR	9.04	8.98	7.62	4.63	8.51	7.96
kNN	TPR	92.9	93.06	94.16	96.14	93.94	94.42
	TNR	94.59	94.85	95.1	96.46	95.94	96.1
	FPR	5.41	5.15	4.9	3.54	4.06	3.9
	FNR	7.1	6.94	5.84	3.86	6.06	5.58

PSO-based feature weighting method achieved the greatest improvements for all machine learning models due to its capability to effectively weight the website features utilized in phishing website detection. In particular, BPNN, SVM, C4.5, RF, and kNN that applied the proposed PSO-based feature weighting achieved the best classification accuracy compared to their performances with applying IG, Chi-square, Wrapper, GA-based features selection, and GA-based features weighting. In addition, the proposed PSO-based feature weighting and GA-based features weighting achieved competitive performances in terms of classification accuracy, and then followed by GA-based features selection for some machine learning models used in this study. Besides, FIGURE 14 shows that the proposed PSO-based feature weighting, Wrapper, GA-based features selection, and GA-based features weighting methods significantly improved the classification accuracy of NB.

In terms of other measures, TABLE 6 shows a comparison of TPR, TNR, FPR, and FNR of the machine learning classifiers improved by the proposed PSO-based feature weighting against their performances with other feature selection and weighting methods.

In terms of TPR, TABLE 6 demonstrates that BPNN, SVM, C4.5, RF, and kNN that improved with the PSO-based feature weighting achieved obviously better TPR compared to these machine learning models with applying IG, Chi-square, Wrapper, GA-based features selection, and GA-based features weighting methods.

That means that BPNN, SVM, C4.5, RF, and kNN improved by the proposed PSO-based feature weighting were able to correctly detect a higher amount of phishing websites.

In terms of TNR, NB, C4.5, RF, and kNN that employed the proposed PSO-based feature weighting outperformed these machine learning models with applying IG, Chi-square, Wrapper, GA-based features selection, and GA-based features weighting methods. Besides, TNRs of BPNN and SVM improved by the proposed PSO-based feature weighting were competitive with TNRs of BPNN and SVM improved by GA-based features selection and GA-based features weighting methods. The higher TNR demonstrates that all machine learning models that employed the proposed PSO-based feature weighting were able to successfully classify a higher number of legitimate websites. Meanwhile, most of the

machine learning models that enhanced with the proposed PSO-based feature weighting achieved FPR and FNR less than FPR and FNR produced by these models after applying IG, Chi-square, Wrapper, GA-based features selection, and GA-based features weighting. The lowest FPR and FNR indicate that the proposed PSO-based feature weighting contributed toward reducing the rate of legitimate websites misclassified as phishing, and decreasing rate of phishing websites misclassified as legitimate, respectively.

In addition to above measures, the average run times (in seconds) of the feature selection and weighting techniques used with the machine learning algorithms were calculated as shown in TABLE 7 using the same computer (PC with processor Intel(R), Core(TM) i7-8550U CPU @ 1.80 GHz 1.99 GHz, and 16 GB RAM). As expected, IG and Chi-square were faster when compared to the other feature selection and weighting techniques. However, IG and Chi-square produced the worst detection accuracy of phishing websites as shown in FIGURE 14 and TABLE 6 since they are filter methods that evaluate the features independently of a specific machine learning algorithm. TABLE 7 also shows that the proposed PSO-based features weighting was faster than GA-based features selection and GA-based features weighting but slower than the Wrapper method for all machine learning algorithms. The GA-based feature selection was the slowest among the feature selection and weighting techniques for all machine learning algorithms.

TABLE 7. The average run times (in seconds) of the feature selection and weighting techniques used with the machine learning algorithms.

	BPNN	SVM	NB	C4.5	RF	kNN
IG	61	90	0.8	1	28	4
Chi-square	59	85	0.5	1	27	6
Wrapper	197	328	1	5	376	194
The proposed PSO-based features weighting	625	974	9	30	694	278
GA-based features selection	2645	5425	26	66	1425	639
GA-based features weighting	1999	4372	23	63	996	630

VIII. CONCLUSION AND FUTURE WORK

In this study, a methodology of the intelligent phishing website detection based on PSO-based feature weighting was suggested. In the proposed PSO-based feature weighting, the website features were weighted with the ideal weights by using PSO to enhance the detection of phishing websites. Consequently, BPNN, SVM, NB, C4.5, RF, and kNN were trained based on the training dataset of features weighted by PSO in order to identify the phishing websites. The experimental results demonstrated that the classification accuracies of BPNN, SVM, NB, C4.5, RF, and kNN were outstandingly enhanced after applying the proposed PSO-based feature weighting. Furthermore, BPNN, SVM, C4.5, RF, and kNN that improved with the PSO-based feature weighting achieved better TPR, TNR, FPR, and FNR compared to the stand-alone machine learning models. This indicated that the machine learning models improved with the proposed PSO-based feature weighting were able to successfully detect and classify both phishing and legitimate websites, respectively. In addition, the proposed PSO-based feature weighting omitted between 7% and 57% of irrelevant features, and only the remaining features were utilized with the classifiers to differentiate the phishing from legitimate websites. Compared to other feature selection methods used in this paper, most of the machine learning models that employed the proposed PSO-based feature weighting outperformed these machine learning models with applying IG, Chi-square, Wrapper, GA-based features selection, and GA-based features weighting. The findings of this paper are expected to influence the future direction of research in phishing website prediction since the feature weighting has not been received much attention by researchers. Furthermore, the machine learning models improved by the proposed PSO-based feature weighting can be used as alternative solutions to effectively detect phishing websites in order to contribute to providing more confidence for customers of online commerce and business. Besides, the use of the most important features set to represent the website can be utilized to speed up the detection process of the phishing website.

The proposed method has demonstrated the benefits of deploying the PSO-based feature weighting in enhancing the intelligent phishing website detection. However, there are a few shortcomings in this study. In the proposed PSO-based feature weighting, we used the original PSO, which utilized the classification accuracy as a fitness objective function during the process of feature weighting. So, the time of feature evaluation and weighting may require a longer time based on the nature of the machine learning algorithm used. Therefore, a faster and improved version of PSO can be used to speed up the performance of the proposed PSO-based feature weighting. Furthermore, other phishing websites datasets should be used to validate and evaluate the proposed PSO-based feature weighting. Lastly, the proposed PSO-based feature weighting was used to enhance some classical machine learning classifiers. Applying the proposed PSO-based feature weighting with ensemble learning and fusion approaches can produce promising solutions with a higher detection accuracy of phishing websites.

ACKNOWLEDGMENT

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (DF-438-830-1441). The authors, therefore, gratefully acknowledge the DSR technical and financial support.

REFERENCES

[1] K. L. Chiew, C. L. Tan, K. Wong, K. S. C. Yong, and W. K. Tiong, "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system," *Inf. Sci.*, vol. 484, pp. 153–166, May 2019.

[2] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, "Machine learning based phishing detection from URLs," *Expert Syst. Appl.*, vol. 117, pp. 345–357, Mar. 2019.

[3] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Tutorial and critical analysis of phishing Websites methods," *Comput. Sci. Rev.*, vol. 17, pp. 1–24, Aug. 2015.

[4] K. L. Chiew, K. S. C. Yong, and C. L. Tan, "A survey of phishing attacks: Their types, vectors and technical approaches," *Expert Syst. Appl.*, vol. 106, pp. 1–20, Sep. 2018.

[5] A. Aleroud and L. Zhou, "Phishing environments, techniques, and countermeasures: A survey," *Comput. Secur.*, vol. 68, pp. 160–196, Jul. 2017.

[6] V. Suganya, "A review on phishing attacks and various anti phishing techniques," *Int. J. Comput. Appl.*, vol. 139, no. 1, pp. 20–23, Apr. 2016.

[7] N. Abdelhamid, A. Ayesh, and F. Thabtah, "Phishing detection based associative classification data mining," *Expert Syst. Appl.*, vol. 41, no. 13, pp. 5948–5959, Oct. 2014.

[8] I. Qabajeh and F. Thabtah, "An experimental study for assessing email classification attributes using feature selection methods," in *Proc. 3rd Int. Conf. Adv. Comput. Sci. Appl. Technol. (ACSAT)*, Dec. 2014, pp. 125–132.

[9] R. S. Rao and A. R. Pais, "Detection of phishing Websites using an efficient feature-based machine learning framework," *Neural Comput. Appl.*, vol. 31, pp. 3851–3873, Jan. 2018.

[10] W. Ali, "Phishing Website detection based on supervised machine learning with wrapper features selection," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 9, pp. 72–78, 2017.

[11] APWG. *Phishing Activity Trends Report 3rd Quarter 2019*. Accessed: Mar. 21, 2020. [Online]. Available: https://docs.apwg.org/reports/apwg_trends_report_q3_2019.pdf

[12] C. Pham, L. A. T. Nguyen, N. H. Tran, E.-N. Huh, and C. S. Hong, "Phishing-aware: A neuro-fuzzy approach for anti-phishing on fog networks," *IEEE Trans. Netw. Service Manage.*, vol. 15, no. 3, pp. 1076–1089, Sep. 2018.

[13] Google. *Google Safe Browsing*. Accessed: Mar. 21, 2020. [Online]. Available: https://safebrowsing.google.com/

[14] *SORBS*. Accessed: Mar. 21, 2020.[Online]. Available: http://www.sorbs.net

[15] *URIBL*. Accessed: Mar. 21, 2020. [Online]. Available: http://uribl.com

[16] *SURBL*. Accessed: Mar. 21, 2020. [Online]. Available: http://www.surbl.org

[17] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Predicting phishing Websites based on self-structuring neural network," *Neural Comput. Appl.*, vol. 25, no. 2, pp. 443–458, Aug. 2014.

[18] M. He, S.-J. Horng, P. Fan, M. K. Khan, R.-S. Run, J.-L. Lai, R.-J. Chen, and A. Sutanto, "An efficient phishing Webpage detector," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12018–12027, Sep. 2011.

[19] H. H. Nguyen and D. T. Nguyen, "Machine learning based phishing Web sites detection," in *AETA 2015: Recent Advances in Electrical Engineering and Related Sciences*. New York, NY, USA: Springer, 2016, pp. 123–131.

[20] S. Y. Yerima and M. K. Alzaylaee, "High accuracy phishing detection based on convolutional neural networks," in *Proc. 3rd Int. Conf. Comput. Appl. Inf. Secur. (ICCAIS)*, Riyadh, Saudi Arabia, Mar. 2020, pp. 19–21.

[21] E. Zhu, Y. Chen, C. Ye, X. Li, and F. Liu, "OFS-NN: An effective phishing Websites detection model based on optimal feature selection and neural network," *IEEE Access*, vol. 7, pp. 73271–73284, 2019.

[22] P. Yang, G. Zhao, and P. Zeng, "Phishing Website detection based on multidimensional features driven by deep learning," *IEEE Access*, vol. 7, pp. 15196–15209, 2019.

[23] M. Kaytan and D. Hanbay, "Effective classification of phishing Web pages based on new rules by using extreme learning machines," *Anatol. J. Comput. Sci.*, vol. 2, no. 1, pp. 15–36, 2017.

[24] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Intelligent rule-based phishing Websites classification," *Inf. Secur., IET*, vol. 8, no. 3, pp. 153–160, May 2014.

[25] P. A. Barraclough, M. A. Hossain, M. A. Tahir, G. Sexton, and N. Aslam, "Intelligent phishing detection and protection scheme for online transactions," *Expert Syst. Appl.*, vol. 40, no. 11, pp. 4697–4706, 2013.

[26] R. M. Mohammad, F. Thabtah, and L. McCluskey, "An assessment of features related to phishing Websites using an automated technique," in *Proc. Int. Conf. Internet Technol. Secured Trans.*, 2012, pp. 492–497.

[27] M. Aydin, I. Butun, K. Bicakci, and N. Baykal, "Using attribute-based feature selection approaches and machine learning algorithms for detecting fraudulent Website URLs," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 774–779.

[28] A. A. Ubing, S. Kamilia, A. Abdullah, N. Jhanjhi, and M. Supramaniam, "Phishing Website detection: An improved accuracy through feature selection and ensemble learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 252–257, 2019.

[29] W. Ali and A. A. Ahmed, "Hybrid intelligent phishing Website prediction using deep neural networks with genetic algorithm-based feature selection and weighting," *IET Inf. Secur.*, vol. 13, no. 6, pp. 659–669, Nov. 2019.

[30] F. Thabtah and N. Abdelhamid, "Deriving correlated sets of Website features for phishing detection: A computational intelligence approach," *J. Inf. Knowl. Manag.*, vol. 15, no. 4, pp. 1–17, 2016.

[31] M. Khonji, A. Jones, and Y. Iraqi, "An empirical evaluation for feature selection methods in phishing email classification," *Comput. Syst. Sci. Eng.*, vol. 28, no. 1, pp. 37–51, 2013.

[32] M. Aburrous, M. A. Hossain, K. Dahal, and F. Thabtah, "Predicting phishing Websites using classification mining techniques with experimental case studies," in *Proc. 7th Int. Conf. Inf. Technol., New Generat. (ITNG)*, 2010, pp. 176–181.

[33] A. Das and S. Das, "Feature weighting and selection with a Pareto-optimal trade-off between relevancy and redundancy," *Pattern Recognit. Lett.*, vol. 88, pp. 12–19, Mar. 2017.

[34] J. Pérez-Rodríguez, A. G. Arroyo-Peña, and N. García-Pedrajas, "Simultaneous instance and feature selection and weighting using evolutionary computation: Proposal and study," *Appl. Soft Comput.*, vol. 37, pp. 416–443, Dec. 2015.

[35] S. Paul and S. Das, "Simultaneous feature selection and weighting—An evolutionary multi-objective optimization approach," *Pattern Recognit. Lett.*, vol. 65, pp. 51–59, Nov. 2015.

[36] L. Jiang, C. Li, S. Wang, and L. Zhang, "Deep feature weighting for naive Bayes and its application to text classification," *Eng. Appl. Artif. Intell.*, vol. 52, pp. 26–39, Jun. 2016.

[37] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Electr. Eng.*, vol. 40, no. 1, pp. 16–28, Jan. 2014.

[38] K. Varpa, K. Iltanen, and M. Juhola, "Genetic algorithm based approach in attribute weighting for a medical data set," *J. Comput. Med.*, vol. 2014, pp. 1–11, Sep. 2014.

[39] B. Xue, M. Zhang, and W. N. Browne, "Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms," *Appl. Soft Comput.*, vol. 18, pp. 261–276, May 2014.

[40] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Nov. 1995, pp. 1942–1948.

[41] A. Kawamura and B. Chakraborty, "A hybrid approach for optimal feature subset selection with evolutionary algorithms," in *Proc. IEEE 8th Int. Conf. Awareness Sci. Technol. (iCAST)*, Jan. 2017, pp. 564–568.

[42] L. M. Abualigah, A. T. Khader, and E. S. Hanandeh, "A new feature selection method to improve the document clustering using particle swarm optimization algorithm," *J. Comput. Sci.*, vol. 25, pp. 456–466, Mar. 2018.

[43] D. O'Neill, A. Lensen, B. Xue, and M. Zhang, "Particle swarm optimisation for feature selection and weighting in high-dimensional clustering," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2018, pp. 1–8.

[44] M.-Y. Cho and T. T. Hoang, "Feature selection and parameters optimization of SVM using particle swarm optimization for fault classification in power distribution systems," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–9, 2017.

[45] J. Wei, Z. Jian-qi, and Z. Xiang, "Face recognition method based on support vector machine and particle swarm optimization," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4390–4393, Apr. 2011.

[46] D. Dua and C. Graff. UCI Machine Learning Repository. School of Information and Computer Science, University of California, Irvine, CA, USA. Accessed: Jan. 10, 2020. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Phishing+Websites

[47] A. K. Jain and B. B. Gupta, "Towards detection of phishing Websites on client-side using machine learning based approach," *Telecommun. Syst.*, vol. 68, no. 4, pp. 687–700, Aug. 2018.

[48] W. Ali, "Hybrid intelligent Android malware detection using evolving support vector machine based on genetic algorithm and particle swarm optimization," *IJCSNS*, vol. 19, no. 9, p. 15, 2018.

WALEED ALI received the B.Sc. degree in computer science from the Faculty of Science, Taiz University, Yemen, in 2005, and the M.Sc. and Ph.D. degrees in computer science from the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia, in 2009 and 2012, respectively. He has been an Associate Professor with the IT Department, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, since September 2013. He has published many papers in several high-impact factor journals, good conferences, and book chapters. His research interests include intelligent web caching, intelligent web prefetching, web usage mining, intelligent phishing website detection, intelligent android malware detection, and machine learning techniques and their applications. In 2012, he received the Best Student Award 2012 from the School of Graduate Studies (SPS), UTM.

SHARAF MALEBARY (Member, IEEE) received the Ph.D. degree in computer science and engineering from the University of South Carolina, USA. He is currently an Assistant Professor and the Head of the IT Department, Faculty of Computing and Information Technology, King Abdulaziz University, Rabigh, Saudi Arabia. He is a very motivated, bright, and social individual. He follows new trending technology in his research. Thus, his publications form a fruitful package of cutting-edge technologies at each year of publication. He prefers applicable research with high impact on the real world though he has done some theoretical work in the beginning of his Ph.D. journey. Although, he is considered a freshly graduate, nevertheless he has proven his ability to do research and be a Successful Leader. As a result, he was nominated to be the IT-Department Head. His research interest includes neural networks to solve research problems in the information technology and computer science fields. His most recent research got published at *ACM-Interactive, Mobile, Wearable and Ubiquitous Technologies*, in which he used machine learning to detect children use of smart devices using a single touch or swipe on the screen.