IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 1, 2020, accepted March 20, 2020, date of publication March 27, 2020, date of current version April 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983219

Identification Failure Data for Cluster Heads Aggregation in WSN Based on Improving Classification of SVM

THI-KIEN DAO[1], TRONG-THE NGUYEN[1,2], JENG-SHYANG PAN[3], YU QIAO[3], AND QUOC-ANH LAI[4]

[1]Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou 350014, China
[2]Department of Information Technology, Haiphong University of Manage and Technology, Haiphong 180000, Vietnam
[3]College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266510, China
[4]Department of Pattern Recognition and Image Processing, Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi 100000, Vietnam

Corresponding authors: Trong-The Nguyen (vnthe@hpu.edu.vn) and Jeng-Shyang Pan (jengshyangpan@gmail.com)

This work was supported in part by the Institute of Information Technology under Project CS20.07, in part by the Senior Mission under Grant NVCC02.01/20-20, and in part by the Vietnam Academy of Science and Technology under Project VAST01.07/19-20.

ABSTRACT Wireless sensor network (WSN) has been paid more attention due to its efficient system of communication devices for transferring information from a target environment to the base station (BS) through wireless links. Precise collecting information from sensor nodes for aggregating data in Cluster Head (CH) is an essential demand for a successful WSN application. This paper proposes a new scheme of identifying collected information correctness for aggregating data in CHs in hierarchical WSN based on improving classification of Support vector machine (SVM). The optimal parameter SVM is implemented by an improved flower pollination algorithm (IFPA) to achieve classification accuracy. The collecting environmental information like temperature, humidity, etc., from sensor nodes to CHs that classify data fault, aggregate, and transfer them to the BS. Compared with some existing methods, the proposed method offers an effective way of forwarding the correct data in WSN applications.

INDEX TERMS Wireless sensor network, support vector machine, identification failure data, classification.

I. INTRODUCTION

Thanks to the rapid development and maturity technologies of wireless communication, sensor, microelectronic, and embedded application, Wireless sensor networks (WSN) have been gradually gaining popularity [1]. WSN refers to a set of spatially dispersed and dedicated sensors for recording and monitoring the physical conditions and organizing the data collected at a central location [2]. WSN often is implemented in distributing infrastructure-freely, having fault-tolerant, scalable, and dynamic [3], [4]; thus, it is widely used for environmental pollution supervision, smart grid, biomedical health management, and behavioral habit detection [5]. The needed ecological information like temperature, light, sound, humidity, wind, air, and water pollution levels could be captured and measured by sensor nodes of WSN [6], [7]. A well designed and employed WSN often follows clustering fashion as efficient ways of saving energy networks that are to generate clusters by arranging the sensor nodes into groups [8]. The power consumption of WSNs is affected directly by the clustering criterion problem [9]. The clustering-based on WSN, such as the hierarchical clustering WSN in which the cluster composes member nodes (MNs) and cluster head (CH). CH is selected from among MNs. The functions of CHs are not only to collect the information from the MN but also to aggregate captured data that forwarded to the BS [2], [10]. FIGURE 1 shows an example of the clustering fashion of WSN.

The style of the clustering-based WSN provides various advantages like energy efficiency, prolonging lifetime, scalability, and less delay. However, the clustered-based WSNs also have got the drawback such as the aggregated data fault problem at CHs that caused the network reliability of the monitoring, and predicting applications decreased [8], [11].

This paper considers the correct data to transfer from CHs to the base station (BS) of WSN. The decision function of the classifier is deployed in the CH for aggregating accuracy data. The cluster heads (CHs) in hierarchical WSN aggregate

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

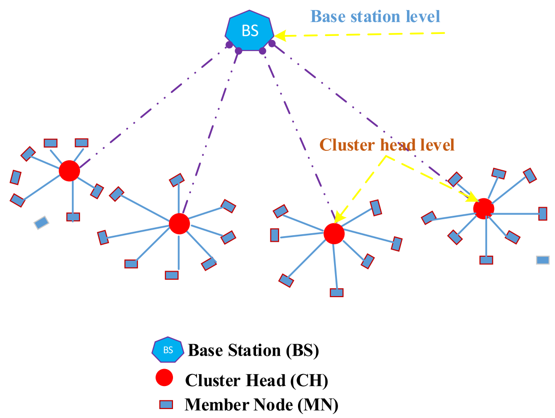


FIGURE 1. An example of the clusters in a wireless sensor network (WSN).

the captured data by MNs, then CHs send them to the BS or via hops (via other CHs). The method of classification is used to detect the faults based on the data learning model in making decisions by combining expert knowledge and statistical learning method. The accuracy of captured data has an essential role in successful ones for several applications such as weather prediction, military monitoring, traffic monitoring, seismic activity prediction, and healthcare monitoring [12].

As deployed network through wireless links of devices on ubiquitous in uncertain and hazardous areas, e.g., battlefields, forest, healthcare, volcanos, highways that often caused the WSN to occur faults commonly [13]. One of the urgent requirements to guarantee proper functioning applications in WSN is reliable data to transfer from CH to the BS for further processing. The distinction between good and faulty data must be determined correctly. The identifying fault data also should be rapid and precise. Identifying the data fault which occurs spontaneously is difficult as those faults may cause continuous failures [14]. The flaws will cause the WSN application to increase data network traffic and wastes battery power [15]. Data fault identification is a promising way to enhance bandwidth, integrity, and reliability. The applied classification is one of the favors of solutions for identifying faults in WSN [16].

This paper considers the collaborating techniques to construct an information identification model for aggregating data in WSNs based on optimizing the classification of Support Vector Machine (SVM) [17] and improving heuristic of algorithm Flower pollination algorithm (FPA) [18]. The optimized parameters for the classifier SVM and its kernel function that are as the factor to validate the information identifying accuracy in WSN applications.

The decision function of the classification technique is deployed in CHs of the hierarchical WSN to classify sensing data from MNs and to identify its faults for the next steps of processing, e.g., aggregating data. It means CHs aggregate sensed data of the environmental information, and transfer them to the BS. Identifying correctness information is figured out by detecting data fault in CHs to enhance the integrity and reliability of the WSN application. The parameters of the classifier SVM used to establish the identification model that is optimized by the proposed improving FPA algorithm (IFPA).

The contributions behinds this paper are as follows.

- Improve Flower pollination algorithm (IFPA) based on enhancing diversity pollens.
- Improve classification by applying IFPA to optimize the parameters for the SVM and kernel functions.
- Identify failure data in Cluster heads in WSN based on the improvement of classification.

The rest of this paper is organized as follows: Section 2 describes related work concludes the fault data issues in WSN, and a recent heuristic algorithm FPA. Section 3 introduces the proposed improvements in FPA (IFPA) and its performance test. Section 4 presents the proposed method of identifying sensed information for CHs in WSN. In Section 5, several experiments are carried on the scenario to evaluate the performance of the proposed method. Finally, the conclusion is discussed in Section 6.

II. RELATED WORK

A. FAULT DATA ISSUES IN WSNs

The electronic components in the sensor nodes of WSN are easy to break-down because of network deployed through wireless links of devices on ubiquitous in uncertain and hazardous areas, e.g., battlefields, forest, healthcare, volcanos, and highways [19]–[21]. The frequent failures happen of WSNs are classified into some types, such as data loss, gain signal, or drift as belonging to hardware, software, and communication failures [22]. Some solutions to the WSN failures were introduced models, such as centralized, distributed, hybrid fault identifications [23]. Data loss or gain caused by hardware failures that occur due to the negligence of sensing capability, power (battery failure) location, and processing units of sensors — for example, the battery failure cause to impairment of sensors. Data loss, gain or drift caused by software failures, e.g., the fusion and aggregation that occurs due to problems in sensor programs. Data loss or drift caused by communication failures occurs the transceiver as the sending and receiving data disrupted from the sensors.

Data faults might occur either separately or simultaneously together and also might happen over a while or instantly. These defects in WSN also can be categorized based on two aspects according to the time of the error and location of the fault. The time-span of the failure indicates the duration of the responsibility. The location of the fault suggests the environment the fault occurs [24]. The period the defects can be categorized into persistent faults and transient faults. Persistent errors are permanent faults that can be solved when the system recovery made. Transient faults are temporary faults that occur due to network congestion or climatic condition. The location of responsibility it is broadly categorized into data-centric and system-centric. Data-centric failure consists of the offset fault, gain fault, drift fault, data loss fault, hard over fault, spike fault, and fusion fault. System centric fault causes data loss or drifts consists of the calibration fault, battery failure, and hardware failure that can lead to the malfunctioning of the entire network.

This subsection reviews several previous works that developed to deal with these mentioned issues of detecting fails data in WSN. A practical method of fault detection in WSN based on the SVM with the Gaussian kernel was deployed for the real-time data classification [24]. In this method, the SVM was used as a classifier of the supervised learning model to classify the fault and proper data. The composition of Lagrangian coefficients and support vectors is to produce the data blocks for a decision function. The goal of the work was to determine the decision function implemented in the CHs to detecting fault data in real-time data classification. The detecting results rate of the work was pretty hight.

Another work used SVM to decide the boundaries of the distribution energy of sensor nodes efficiently in WSN [25]. The training of SVM based on distributed incremental learning in a WSN was analyzed the performance of the network life. As the same managing energy in WSN, the SVM was also used as a classifier to identify the level powers of nodes fro prolonging network life [26]. A study of classifying vehicle targets [27] in WSNs was implemented by applying SVM for cross-matching or voting on the car classification through sensor nodes.

Though the methods got good results of classification, they had the disadvantage, e.g., not be optimized, nor flexible applied with the environment changing. The reasons for existing drawbacks are the set parameters for SVM and kernel functions not yet optimized, but only manually with the fixed practice values.

As the same style of the solution to the intrusion or fault diagnosis or error detection in datasets WSN, a hidden Markov model (HMM) [28] was used as the centralized technique to identify offset fault, gain fault, and struck fault [29], [30]. The method of the statistical model HMM applied to process with unknown variables to determine the hidden parameters from the observable parameters. Each state has a possible output distributed probability by HMM generating sequence of state. The state is not able directly observed by HMM, but its variables influenced by the state are observable. The hidden states between sensor nodes determined the health conditions of WSN by applying the HMM model to estimate faulty data situation probability [31]. An approach detection errors in big data sets WSN was developed based on the cloud computing scheme [32]. The cloud-based technique is the hybrid-based fault detection technique, where the data collected from sensors are stored in cloud storage.

The decision tree (DT) was also introduced as a greedy divide-and-conquer algorithm to detect intrusion in WSN [33]. The classification air quality on the WSN monitoring system was implemented based on the applied DT algorithm [34].

The tree is constructed in a top-down recursive manner as assumed attributes are categorical with continuous traits also can be handled. All attributes for a given node belong to the same class. At the start, all the training examples are at the root, and then instances are partitioned recursively based on selected attributes. Attributes are selected based on an impurity function (e.g., information gain). Conditions for stopping partitioning consist of one of the requirements as follows. There are no remaining attributes for further partitioning – the majority class is the leaf, and there are no examples left.

Moreover, the neural network (NN) was applied to identify fault detection in WSNs [35]. A two-stage neural network was used to classify various bearing faults to correct temperature error in WSNs [36]. The NN estimated the principal components and used a supervised learning vector quantization network and a self-organizing map scheme in WSN.

Moreover, the decision tree learning algorithm is a greedy divide-and-conquer algorithm. Intrusion detection in WSN has introduced through decision trees (DT). Assume attributes are categorical now (continuous traits also can be handled). The tree is constructed in a top-down recursive manner. At the start, all the training examples are at the root. Instances are partitioned recursively based on selected attributes. Attributes are selected based on an impurity function (e.g., information gain). Conditions for stopping partitioning consist of one of the conditions as follows. All attributes for a given node belong to the same class. There are no remaining attributes for further partitioning – the majority class is the leaf, and there are no examples left.

The practical solutions for identifying data, intrusion, and diagnosis detection schemes in WSN can be found in [22] [37], [38]. Those methods were lack of optimization and flexible implementation in an environment changing.

B. FLOWER POLLINATION ALGORITHM

Flower pollination algorithm (FPA) is one of the most popular recent heuristic algorithms [18]. The inspiration for designing FPA is taken from the pollinating process flower plants. Two processes of pollination flowers in FPA applied to formulate the optimization equations, such as self-pollination and cross-pollination. Self-pollination is a process of pollinating flowers that is taken place in the same plants. The cross-pollination is the pollinating flower in different plants. The transporting pollen of flowers has to obey the rules of Lévy flight. The self-pollination of flowering plants is regarded as the local pollination in the searching area that is the concept conventional called exploration in the population-based algorithm. The pollination of flowers, i.e., pollens in the same plant or the flowers, guides the optimization process. So, the local pollination is modeled as follows:

$$x_i^{t+1} = x_i^t + u \times (x_j^t - x_k^t) \tag{1}$$

where x_j^t and x_k^t are vectors of the pollens as solutions in optimization algorithm; u is a random number that is drawn from a uniform distribution in [0, 1]. If x_j^t and x_k^t come from the same plans or the same selected population, u would become a local random walk. The process of pollination can be expressed by the equation of updating the solution vector. In the process of pollen transmission, insects may carry pollen as carriers. This processing of pollination, the cross-pollination is considered as a global pollination search exploring in a promising area. Updating the location of global pollination is simulated as follows.

$$x_i^{t+1} = x_i^t + \gamma \times L(\lambda) \times (x_i^t - g^*) \tag{2}$$

where t is the current iteration ; g^* is the current best solution found so far; γ is a scaling factor to control the step size. Steps of different lengths are used to represent the Levy flight of insects caused by long-distance flight. The Lévy distribution is as follows

$$L = \frac{\lambda \Gamma(\lambda) \times \sin(\frac{\pi\lambda}{2})}{\pi \times s^{i+\lambda}} \tag{3}$$

where $\Gamma(\lambda)$ is the gamma function, and its valid distribution is the step with s > 0. The step size, like a parameter L(λ) corresponds to the strength of the pollination. A parameter (notation is $p \in [0, 1]$) is used as the approach probability to switch between global and local pollination as Eq.(1) and (2) respectively in the process of flower solution. To improve the effectiveness of the identification scheme suitable to the limitation of resources of WSN, we propose a new improving version of the FPA for optimizing the SVM kernel parameters to establish the identifying collected information scheme in CHs of WSNs. Improving FPA (IFPA) is presented in the next section.

III. IMPROVED FLOWER POLLINATION ALGORITHM

A. THE PROPOSED IFPA ALGORITHM DESCRIPTION

This section presents an improved Flower Pollination Algorithm (IFPA) based on a novel communication strategy. We introduce a parallel operation by dividing the population into some groups to enhance the global search capability of the algorithm. A new communication strategy is proposed based on the ideal of high-quality pollens of each group are combined for evolution and replaced the old pollens. The evolution of species inspired the approach naturally, which refers to the mutation, crossover, and selection operation in Differential evolution [39]. The policy can take the advantage to make the better pollens into improved through mutation and crossover; thus, the algorithm may be easy to jump off the local optimum. This strategy collected the pollens with functional fitness in each group to form a new population for evolving and replacing the original pollens. The detail steps of the IFPA method are presented as follows.

Step 1. For every R_1 iterations (R_1 is a number of iterations)) pollens of each group are strictly sorted by their fitness from worst to best.

Step 2. Combine each group's pollens with the best first quarter fitness into a new population x^t for evolution.

Step 3. The new population is generated with mutating according to the following formula.

$$z_i^{t+1} = x_{r_1}^t + F \times (x_{r_2}^t - x_{r_3}^t) \tag{4}$$

where z_i^{t+1} is the ith mutated pollen in $(t+1)$th iteration, $x_{r_1}^t, x_{r_2}^t, x_{r_3}^t$ are the different random pollens in the tth generation of the new population. F is an adaptive coefficient of mutation [40], and $F \in [0, 2]$. It is calculated by the following formula.

$$F = F_0 \times \left(2^{\alpha}\right), \quad \alpha = e^{1-\frac{t_{max}}{t_{max}+(1-t)}} \tag{5}$$

where F_0 is mutation constant, t is present iteration, $t_{\max}$ is the maximum iteration.

Step 4. The test pollen is obtained by crossover operation.

$$u_{i,d}^{t+1} = \begin{cases} z_{i,d}^{t+1}, & \text{if } (d = d_{rand}) \text{ or } (rand(0,1) \le CR) \\ x_{i,d}^{t+1}, & \text{otherwise} \end{cases} \tag{6}$$

where CR is cross-over probability, i is the ith pollen, $i = 1, .., N_P$, in which N_P is the population's size, d is the dimension of pollen, d_{rand} is a randomly selected number of sequences $[1, \ldots, D]$, and D is max pollen's dimension.

Step 5. If the fitness value of the test pollen u_i^{t+1} is better than the contemporary pollen x_i^{t+1}, $x_i^{t+1} = u_i^{t+1}$, otherwise $x_i^{t+1} = x_i^t$ that is as the following description.

$$x_i^{t+1} = \begin{cases} u_i^{t+1}, & \text{if}(f\left(u_i^{t+1}\right) \le f\left(x_i^t\right)) \\ x_i^t, & \text{otherwise} \end{cases} \tag{7}$$

Step 6. The old group pollens x^t is replaced one by one by the new evolution pollens x^{t+1}, achieving communication between groups. The steps of the scheme repeat until it meets the termination condition.

B. EXPERIMENTAL RESULTS OF TESTING IFPA

For meeting test functions' diversity, the eight benchmark functions from the CEC2013 [41] are utilized to test the performance of the proposed IFPA, the detail presetting iteration, and boundaries are listed in Table 1. In order to make the following results comparable, IFPA is compared with the original FPA, DE, and PSO, respectively. The population size of the four algorithms (IFPA, FPA, DE, and PSO) in comparison is the same that N_p set to 80. The switch probability p is set to 0. 6. The parameter R_1 equals to 10, as communication should be done every ten iterations, and the mutation constant F_0 equals to 0.5, crossover constant CR equals to 0.5, a number of the groups G_n set to 4. Because of the strong randomness of the meta-heuristic algorithm, the number of runs is set to 20.

Table 2 summarizes the comparison of the average obtained results of 20 runs of the proposed IFPA with the FPA, PSO, and DE algorithms. In Table 2, a parameter of rate difference symboled r is used as a comparison ratio that is a pairwise comparison between the IFPA and FPA, DE, and PSO algorithms, respectively. Three symbols, "+," "−," and "∼," respectively, are used to indicate that the comparison result is the "better," "worse," and "similar." For example, if the IFPA method is better than the FPA method, the result is set to the sign "+." For the results are the worse, and similar symptoms are treated the doing same with "−," and "∼" respectively.

It is observed from Table 2 that among the results of the test benchmark function, the number symbol of '+'s is more

TABLE 1. Eight selected benchmark functions.

No.	Test Functions	Range	Dimension	Iteration
1	$f_1(x) = \sum_{i=1}^{N} x_i^2$	±100	30	2000
2	$f_2(x) = \sum_{i=1}^{D} (10^6)^{\frac{i-1}{D-1}} x_i^2$	±100	30	2000
3	$f_3(x) = \sum_{i=1}^{D} [10 + x_i^2 - 10cos2\pi x_i$	±100	30	2000
4	$f_4(x) = 418.983n - \sum_{i=1}^{N} x_i \times sin\left(\sqrt{\lvert x_i \rvert}\right)$	±100	30	2000
5	$f_5(x) = 418.983n - \sum_{i=1}^{N} x_i \times \sin\left(\sqrt{\lvert x_i \rvert}\right) + f_5^*$	±100	30	2000
6	$f_6(x) = \sum_{i=1}^{n} \sum_{k=1}^{i} x_i$	±10	30	2000
7	$f_7(x) = random[0,1) + \sum_{i=1}^{N} i \times x_i^4$	±10	30	2000
8	$f_8(x) = \sum_{i=1}^{n-1} (100 \times (x_{i-1} - x_i^2)^2 + (1 - x_i)^2$	±600	30	2000

than the symbols in the table that mean the IFPA has won in the competition in the comparison. The results of IFPA are almost better than other competitive methods. As a result, it shows that the proposed method has better performance than the original algorithm.

FIGURE 2 shows the compared curves of the obtained results of the proposed IFPA with the other algorithms, e.g., FPA, PSO, and DE. Observed from FIGURE 2, the results of the proposed IFPA method are almost better than other competitive methods in terms of the converge rate of the optimization process. As a result, it shows that the proposed method has better performance than the original algorithm.

IV. IDENTIFYING SENSED INFORMATION SCHEME

The proposed scheme aims to establish a decision function in the CHs that can be used in real-time for aggregating the precise data that classified any new collecting data from MNs and forwarding them to the base station in WSN. An elaborate scheme is working on CHs that consumes a lot of energy. However, some WSN applications can have CHs that equipped with unlimited power, and the data learning model in making decisions for a decision function is figured out by combining expert knowledge and optimal classification. The decision function of the classifier is deployed lightly in the CH for aggregating accuracy data. The proposed scheme of identifying sensed information consists of the steps of collecting data, pre-processing data, optimizing parameters, detecting faults, and classifying data.''

The collected data from various sources by the sensor nodes for a system input could consist of noise or faults caused by different reasons that lead to misclassification data. The data should be 'cleaned' data by removing the noise before further processing data. The detail of the steps of the designing scheme is presented as following subsections.

A. COLLECTING AND PREPROCESSING DATA

Collecting data from sensor nodes may consist of the redundancy of the attributes that cause the misidentification. Attribute selection is crucial to remove the excess among the features of time-series collected data to enhance optimal computation and communication. A subset of attributes is evaluated and compared to other combinations that its characteristics are passed forward and backward to add or remove attributes. A form of collecting data is a vector of the data-points. Let S_i be a value at position i^{-th} with k is kernel size. The full width at half maximum is used as an expression of the extent of function given by the difference between two extreme values of the independent variables at vector data is equal to half of its maximum value.

$$F(\sigma, \mu_i) = \frac{1}{2}\pi\sigma^2 \times exp^{-\frac{1}{2}\left(\frac{\mu_i}{\sigma}\right)^2} \tag{8}$$

where σ is the width of the curve and μ_i is the distance from the origin point i in the horizontal axis. If a current spike occurs in collecting data, that could be the noise value, and the smoothing cost will be calculated and substituted them [42]. A weighted average of smoothed value is for the new points that are calculated as follows.

$$S_i = \begin{cases} \frac{1}{2k+1} \sum_{j=-k}^{k} \mu_{i+j}, & \text{if (no spike)} \\ \sum_{j=-k}^{k} (w_j * \mu_{i+j}), & \text{otherwise} \end{cases} \tag{9}$$

where w_j is the weight factor of smoothing. The advantage of the weighted average is that considered for a straightforward implementation that is for every attribute in the resulting data, and fewer values in the weighted sum. There are some types of faults based on the data-centric in WSN for sensor readings, such as due to hardware failure, software failure, and communication failure [43].

B. TRAINING AND TESTING DATA

There are two phases of the fault detecting data in this section the data learning and real-time decision-making stages. In the data learning phase, the essential elements of the data are respected and maintained in the process. The data learning stage also uses a statistical learning technique. The needed experience from the expertise to resolve different problems

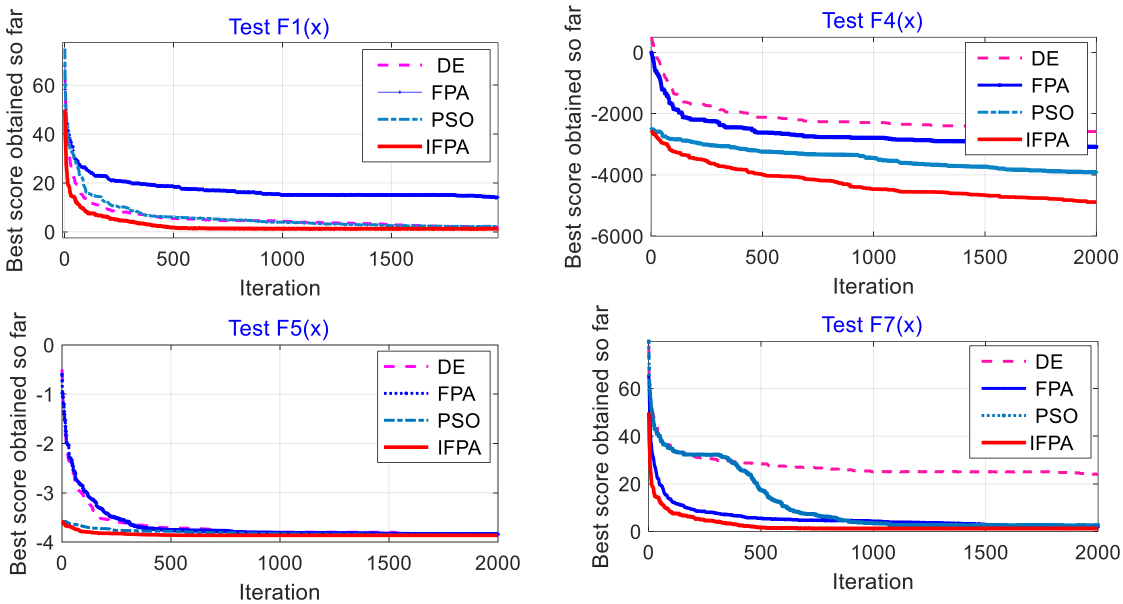


FIGURE 2. Compared curves of the obtained results of the proposed IFPA with the FPA, PSO, and DE algorithms.

TABLE 2. Comparison of the obtained results of the proposed IFPA with the FPA, PSO, and DE algorithms.

Test Functions	Proposed IFPA	FPA	r	PSO	r	DE	r
1	1.50E-323	1.55E-323	+	5.06E-07	+	1.25E-05	+
2	1.40E-322	2.5E-322	+	1.24E-322	-	4.681E-10	+
3	8.40E-322	2.00E-323	-	2.00E-323	-	9.62E-07	+
4	-2,50E-323	-3.90E-323	+	-3.10E-131	+	-1.83E-43	+
5	-4.08E-324	-4.08E-324	~	-4.96E-324	-	-4.92E-324	-
6	2.5E-323	4.8E-323	+	2.35E-13	+	4.43E-10	+
7	7.00E-05	7.61E-05	-	8.11E-05	+	8.15E-05	+
8	2,15E+01	1.96E+01	-	2.32E+01	+	1.96E+01	-
Summary	4+ 1~ 3-			5+ 0~ 2-		6+ 0~ 2-	

affecting WSNs is applied in learning data phrase because the classification based on data learning allows using knowledge in making decisions. The classifiers and the decision function are implemented in the CHs in WSN for aggregating accuracy data. The aim is to establish a real-time decision function in CHs to classify failure data of any new collecting data from sensor nodes MNs. A labeled dataset is used as a learning data composed of a set of regular data and a set of erroneous data. FIGURE 3 shows a model of classifying attributes that have two phases in the model of organizing characteristics.

In the decision-making phase, a new data vector is constructed with blocks of data measurements V_t that included

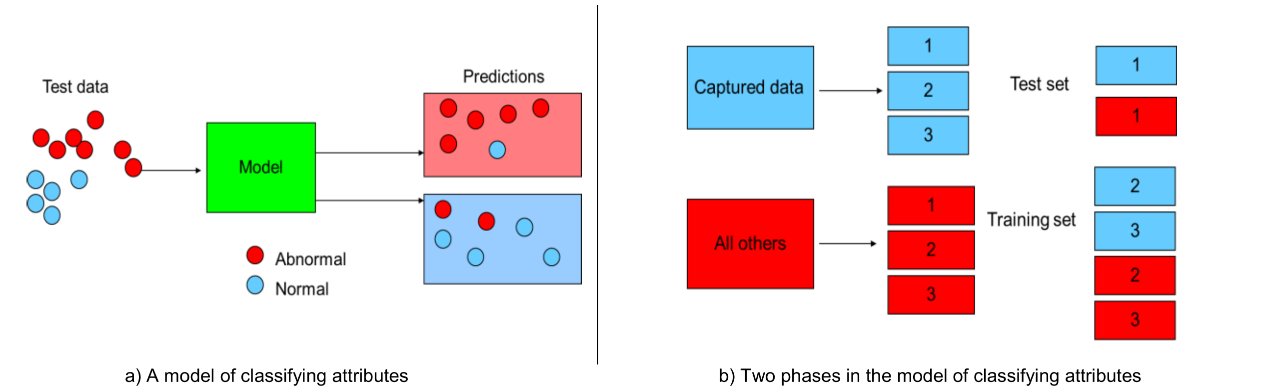


FIGURE 3. A model of classifying attributes that have two phases in the model of categorizing traits.

TABLE 3. Four kernel functions.

No.	Name	Kernel function expression
1	Radial basis function (RBF)	$K(x_i, x_j) = \exp\{-\frac{\|x_i - x_j\|^2}{\sigma^2}\}$, where σ^2 is the radial parameter of basis kernel width
2	Polynomial	$K(x_i, x_j) = ((x_i, x_j) + 1)^\mu$ μ is the distance from the origin point i in the horizontal axis
3	Spline	$K(x_i, x_j) = 1 + (x_i, x_j) + \frac{1}{2}(x_i, x_j) \times \min(x_i, x_j) - \frac{1}{6}min(x_i, x_j)^3$
4	Ensemble RBF (ERBF)	$K(x_i, x_j) = \exp\{-\frac{\|x_i - x_j\|}{\sigma^2}\}$

three measures (V_t, V_{t-1}, V_{t-2}). The decision function will makeover the new data vector. If its result is positive, it belongs to the standard case (a class of normal data functionality); otherwise, it is considered as a faulty case. The computationally inexpensive by applying a simple decision function used in the CH that makes the proposed scheme very efficient with sensors as limited resource nodes. A part of collecting data with the selected attributes would be used to train the captured data by applying the model learning, the kernel function, the cross-validation, and expert's experience values. After the collected dataset is trained, the classifier model would be applied to test data for classification with failure data identification.

The relationship between the size of the training set and feature space, the SVM's performance depends on the optimal parameter of SVM and kernel functions. Table 3 lists some kernel functions.

C. FITNESS FUNCTION FOR OPTIMIZATION CLASSIFICATION

The purpose of identifying correct information is to eliminate failures or redundant data from the collected data. This designated task is to figure out a subset of abnormal attributes from all existing characteristics in collecting data to reaches the highest possible accuracy using that detecting unusual behavior. The number of attributes selected to achieve maximum efficiency is modeled mathematically for considering as evaluation measure for the fitness function. The number of elements chosen or dimension is reduced in the big datasets to increase the classification performance of detecting scheme. Let TA be the total number of attributes contained in collecting data. A new evaluation measure of the objective function is proposed to optimize parameters SVM by evaluating of optimization method as follows.

$$Fitness = \frac{1}{TA \times (-\sigma^2)} \left(\sum_{i=1}^{TA} (x_i - \hat{y}_i)^2 \right) \tag{10}$$

where y_i is the predicted value, x_i is the actual value, and σ is calculated as follows.

$$\sigma^2 = \frac{1}{TA} \sum_{i=1}^{TA} (x_i - MZ)^2 \tag{11}$$

where MZ is mean of accuracy that is computed as follows.

$$MZ = \frac{1}{k} \sum_{i=1}^{k} \partial_i \tag{12}$$

where k is the total number selected attributes; ∂ is the accuracy rate that is calculated as the evaluated measure of the performance of each fold data during the training data. The formula of the accuracy rate ∂ is modeled as follows.

$$\partial = \frac{TP + TN}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{13}$$

where TP is noted as the total number of positive cases that are correctly identified as positive; TN is indicated as the total number of negative instances that are correctly identified as unfavorable; FP is the total number of failures as negative cases that are incorrectly identified as positive, and FN is the total number of failures as positive cases that are incorrectly identified as unfavorable.

D. MODEL OF FAILURE DATA IDENTIFICATION

In essence, the purpose of correctness information identification is to distinguish correctness information through a filter of decision function installed in CH for aggregating data that forward to BS in WSN. Therefore, data identification is a typical classification problem. SVM is one of the most typical classification approaches, has the characteristics: the principle of structural risk minimization is used to solve classification problems. The SVM classifier is adopted to assess classification accuracy in using the selected subset of attributes of collecting data. The typical classification problem would be returned to optimization problems by applying the metaheuristic algorithm with the SVM classifier. The subsets of selected attributes are input into the SVM, and evaluations are carried out using the k-fold cross-validation approach. In the case of identification data for WSN, the cross-validation is performed by the value of k is set to 2 for attributes data with cases below 1000, and k is set to 5 for data attributes with examples above 1000. These k-fold settings were chosen to allow the IFPA to have sufficient facts for the training and testing process.

After obtaining minor errors from training samples, the generalization ability of the learning mechanism can be improved to ensure small errors for test sets. The classification performance of SVM is mainly influenced by suitable setting SVM (C, δ) and kernel function parameters $(\sigma,\ \mu)$. Therefore, it is imperative to study the selection of option SVM parameters and kernel function parameters. To find the suitable values setting SVM and kernel function parameters, the IFPA is applied to find out the best one for specific situations, e.g., the limited energy and computation of nodes in WSN. The penalty factors (C, δ) and kernel parameters σ, μ) of SVM is to set the location x of pollen solution of the IFPA, x is mapped to (C, δ, σ,μ) .

The pollen mapped to the solution is constructed based on the SVM mapping procedure. This procedure is namely SVCoveter(*Inputs, Target, pollen*), in which *Inputs,* and *Target* are the training inputs, and the training target, respectively that obtained from the collected data, and pollen is considered to the searching agent in IFPA. In this situation, structured pollen consists of attributes, e.g., design variables: *pollen.x*, test variable: *pollen.u*, fitness value: *pollen.fit*, accuracy variable: *pollen.acc*, selected variable: *pollen.select*, the number of the feature: *pollen.nfeat*, global best: *pollen.g*, etc.

The predicted vector $\hat{y}$ in Eq.(10) is a probability approximation function denoted $f(y_i|\theta_i)$, in which θ_i is the solution vector obtained by using the data $\{(x_i, y_i)\}$ for training. The θ_i is assigned to each combination of (C, δ, σ,μ). Thus, the design variables: $pollen.x(x_1, x_2, x_3, and x_4)$ is responding to $F(C, \delta, \sigma,\mu)$. The design variables x should be adopted conveniently with the logarithm, e.g., $x_1 = \log_{10} C$, $x_2 = \log_{10} \delta$, $x_3 = \log_{10} \sigma$, and $x_4 = \log_{10} \mu$, with the boundaries that initialized search ranges such as $zx_1 \in [0, 08]$, $x_2 \in [-09, 0]$, $x_3 \in [-2, 01]$, *and* $x_4 \in [-3, 2]$. *pollen.x, u* are updated according to Eqs. (1), (2) and (4), (6) for FPA and IFPA respectively; *pollen.fit* is updated according to Eq.(10); pollen.acc is updated according to Eq(11).

It means that pollen is encoded to its solution. The best pollen solution is obtained optimal pollination according to the fitness function evaluation in Eq.(10) that is the optimal values of the parameters of the SVM. The process of applied IFPA for failure data identification for the CHs of cluster WSN is shown in FIGURE 4

The implementation process of constructing the failure data identification model shown in FIGURE 4, that notes further as follows. In initialization, the pollination position is encoded pollen. N_p is the total number of pollination. In calculation objective function, Fitness f is computed the fitness value of pollen fitness according to the detecting accuracy of SVM, the best g^* is obtained, and allocates pollen to map them respectively. Repeat the operations of generating new local solutions, comparing the evaluation results, finding the best solution, until it meets the termination conditions. Finally, decode the optimal g^*, and then obtain the optimal parameter combination (C, δ, σ,μ) for the classifier to identify failure data.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

Different scenarios of the WSN are simulated, e.g., some set faulty nodes within the network, i.e., increasing or decreasing temperature, humidity. Nodes in WSN sensed data that damaged in traffic congestion conditions. Under each setting conditions, no-fault network and different numbers of faulty nodes within the system are generated for simulation. For an N-node topology, there are some combinations of faulty nodes with varying from one to five, respectively. The information of the packets such as data transmission time and source node IDs received by the CHs are aggregated and forwarded to the BS. Data were randomly picked up by member nodes and were forwarded to the CHs. Then, the CHs will pack the data to be packets and transmit them towards the BS node. There are about 400 data packets under each scenario generated by sensor nodes randomly.

A. ENVIRONMENT SETTING TO COLLECT DATA

Assumed an N-node topology of cluster-based WSN was deployed in scattering the area of $M \times M$ randomly, where N is a number of nodes can be set 100, 200, 300; M is a length measurement of the deployed area that can be 200, 300, 400 meters. The network has a base station (BS) that operates with an unlimited power supply. BS receives the aggregated data from CHs.

The simulation for the proposed method is constructed in Matlab and referenced with the real sensor nodes kits in our Lab. All runs are executed in Matlab 2018b on a Windows 10 Home 64 operating system on a Lenovo T470p laptop with Intel® Core$^{\text{TM}}$ i7-8665U with 8Gb of RAM.

The characteristics of WSN operation was assumed that behaving like scheduling periods of packet transmission time. The proposed method namely IFPA-SVM is compared with

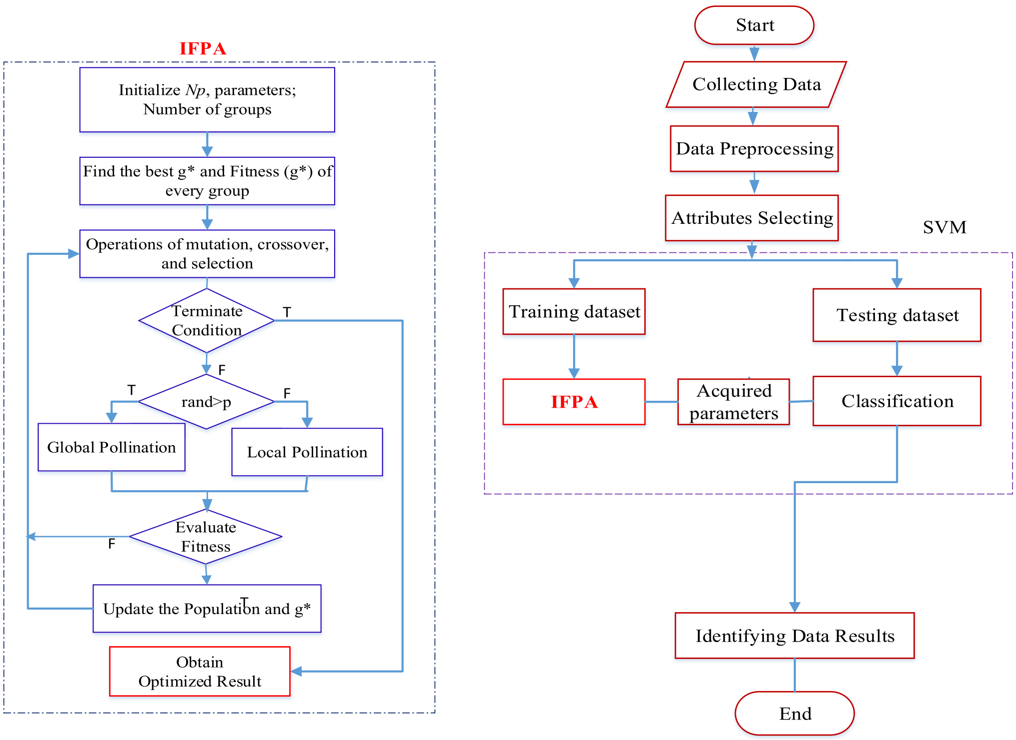


FIGURE 4. The process of applied IFPA for failure data identification in the CHs of cluster WSN.

the other techniques, e.g., GA-SVM [44], PSO-SVM [45], FPA-SVM [46] methods. Table 4 lists the initial values for setting the parameters of the experiment.

The collecting data consists of attributes, e.g., the packet-ID, the node ID, ClusterHeads(e.g., 0 or 1), the sensing data(e.g., temperature, humidity, lights, network status, gas..), radio signal strength (Decibels (dB)), and noisy estimated. The types of sensing data depend on the kind of sensor of the node at the purpose of the application.

The data is split into two parts: one for training and the other for testing with a variety size 70% and 30%, respectively.

The attributes, e.g., temperature, network status, could be 'normal' or 'faulty.' For an example of the network status, if the attribute is class labeled 'normal,' the network consists of no defective sensor. Otherwise, the system includes at least one faulty sensor in both the training and testing phase.

The collecting dataset is taken from outdoor data collection from the clustering-based sensor network with a total of 4001 samples. This hierarchical WSN set a schedule for collecting data of sensing data is a measurement of temperature in a period of every five minutes scheduling loop for packets data with a total of 55 minutes.

B. RESULTS WITH PERFORMANCE OPTIMIZATIONS

We can obtain the optimal results by applying the proposed IFPA for the objective function in Eq.(18). We compare the results of the best achieved so far from the optimizations that

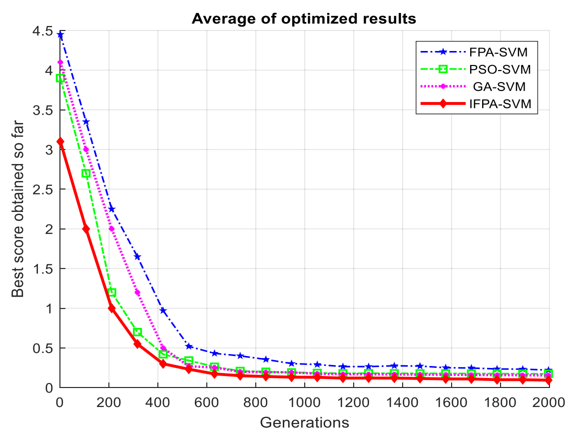
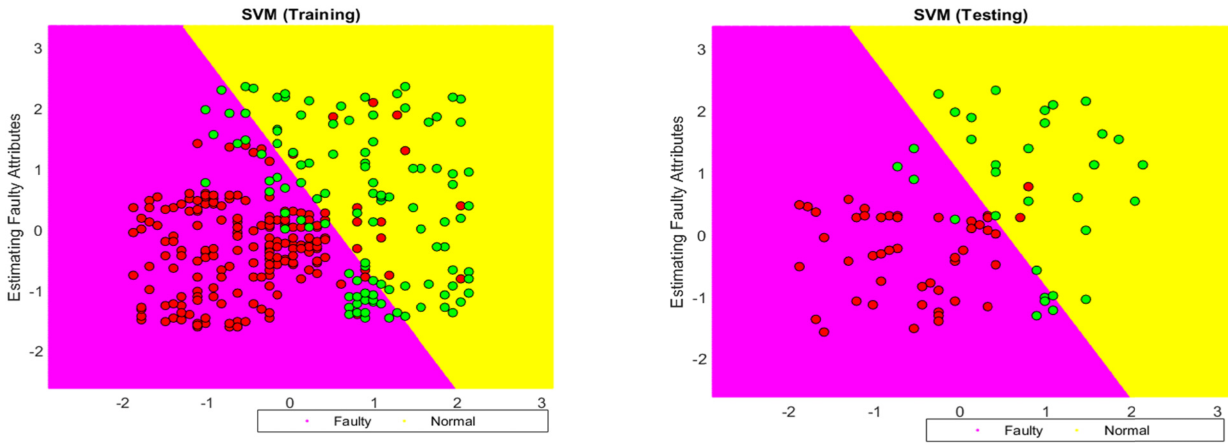


FIGURE 5. Comparison average obtained results of the proposed IFPA-SVM with the FPA-SVM, GA-SVM, and PSO-SVM methods for the objective function.

are averaged with 25 runs over 2000 generations. FIGURE 5 shows the comparison of the proposed IFPA-SVM with the original algorithm FPA, and the other techniques, e.g., GA-SVM, and PSO-SVM methods.

It can be seen from observed FIGURE 5 the curve line of the IFPA is better than the original FPA, and the other techniques of GA-SVM, and PSO-SVM in terms of the converge rate for competitive purposes.

Table 5 summarizes the obtained parameters results of the optimization of processes IFPA, FPA, PSO, and GA with SVM and Kernel functions for the WSN's collected data.



a) A visualization result of the training set

b) A visualization result of the testing set

FIGURE 6. Visualization results of the training and testing set.

TABLE 4. Initial values for setting parameters of the experiment.

Parameters noticed	Denoted symbols	Initial values
Initial node energy	E_j	$0.5J$
Data aggregation energy	E_{DA}	5nJ/bit/signal
Receiving and transmitting energy	E_{fs}	10pJ/bit/m^2
Radio electronics energy	E_{elec}	$50nJ/bit$
Number bit of a data message	l	1024bit
Amplifier energy	E_{mp}	0.013pJ/bit/m^4
Number of nodes in WSN	N	100/200/300/ nodes
Space distribution	M	$100/200$ $/300m$
Generations	$MaxIter$	2000
Number of runs	$runs$	25
GA-SVM	Initialize parameters	$N_P = 80,$ $F_c = 0.8,$ $P_m = 0.1$
PSO-SVM	Initialize parameters	$N_P = 80,$ $c_1, c_2 = 2,$ $w = 0.9$
FPA, IFPA-SVM	Initialize parameters	$N_P = 80,$ $R_{1,} = 10,$ $R_c = rand,$ $F_c = 2$

Moreover, to evaluate the performance diversity of the proposed approach, parameters of statistical variables are used to measure the performance of the proposed model, namely mean-square error (*MSE*) and scatter index (*SI*), which are defined as the following expression.

$$MSE = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N}(O_i - P_i)^2} \times 100\% \tag{14}$$

$$SI = \frac{MSE}{\bar{O}} \tag{15}$$

where N is the number attribute selected of data; O_i and P_i are observed and predicted failures collected data respectively; $\bar{O}$ indicated as average observed failures received information. Table 6 depicts the comparison of statistical measures of the proposed method-SVM model with the other hybrid-SVM methods for different kernel functions in training and testing WSN collected data. Observed the results from Tables 5 and 6, the compared optimal result shows that the proposed IFPA-SVM provides competitive results.

C. RESULTS WITH CLASSIFICATION OPTIMIZATIONS

For validation of the classification accuracy of the proposed method of hybrid optimization with classification, we use two metrics of the identification accuracy (*IA*) and the false positive rate (*FPR*) that are formulated as follows.

$$\mathrm{IA} = \frac{number\ of\ faulty\ data\ identified}{total\ number\ of\ faulty\ data\ present} \tag{16}$$

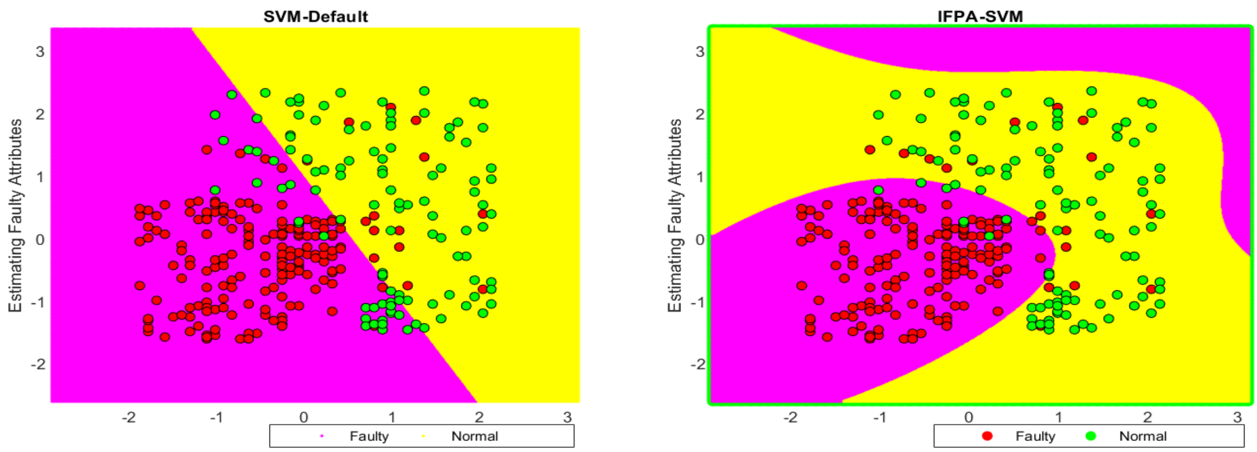
where *IA* is detection accuracy that is clarified as the ratio of the number of erroneous data identified to the total number of current incorrect data.

$$FPR = \frac{number\ of\ non\,faulty\ data\ identified\ as\,faulty}{total\ number\ of\ faulty\ data} \tag{17}$$

Here *FPR* is the false positive rate that is classified as the ratio of some non-faulty data identified as faulty to the total number of incomplete data. Table 7 summarizes the results of the training accuracy of various fault types.

TABLE 5. The obtained parameters results of the optimization of methods IFPA, FPA, PSO, and GA with SVM and kernel functions.

Approaches	Parameters	RBF	Poly-nomial	Spline	ERBF
IFPA-SVM	C	411	131	2	29
	δ	0.000481	0.00035	0.00004	0.000028
	σ	1.12	-	-	5.2
	μ		3	3	-
FPA-SVM	C	401	136	2	24
	δ	0.000481	0.00035	0.00002	0.000021
	σ	1.13	-	-	5.1
	μ		3	3	-
PSO-SVM	C	403	135	2	28
	δ	0.000481	0.00035	0.00003	0.000022
	σ	1.15	-	-	5.5
	μ		3	3	-
GA-SVM	C	402	132	2	24
	δ	0.000481	0.00035	0.00002	0.000023
	σ	1.12	-	-	4.1
	μ		3	3	-
SVM-Defualt	C	298	400	1	20
	δ	0.00031	0.00423	0.00001	0.00002
	σ	1	-	6	-
	μ	-	3	-	-



a) A shot of visualization result of the testing SVM-default

b) A shot of visualization result of the testing IFPA-SVM

FIGURE 7. Visualization results of the testing for SVM and IFPA-SVM.

FIGURE 6 shows the visualization results of the training and testing set. FIGURE 7 depicts the comparison of visualization results of the proposed IFPA-SVM with the SVM default for testing collected data from the WSN. Observed obviously, the proposed method of classification produces more accurate classification then SVM.

The experimental results of the proposed scheme for the classification dataset of collecting temperature are compared with the cases of the methods, e.g., the support vector machine (SVM) [24], the decision tree (DT) [33], and the hidden Markov model (HMM) [30] concerning the false positive rate of the training set. FIGURE 8 shows a comparison of the proposed method (IFPA for the optimal parameters of SVM and kernel function, namely IFPA-SVM) according to the FPR metric with the original FPA, SVM, DT, and HMM. The proposed method provides the best value of FPR that helps to make a significant and essential improvement of FPR as compared to others. The growth is starting from 55% compared to SVM, and reaches 65%, 69% compared to HMM, DT, respectively.

In the measurement of the results of the proposed scheme in comparison with other methods, a mathematical tool

TABLE 6. Comparison of statistical measures of the methods with the SVM model for different kernel functions.

Methods	Processing	Measures	RBF	Poly-nomial	Spline	ERBF
IFPA-SVM	Training Data	*MSE*	3.0159	2.1841	2.0835	4.0305
		SI	0.2758	0.2178	0.1838	0.3349
	Testing Data	*MSE*	3.1069	2.2812	3.2355	4.5208
		SI	0.2157	0.2145	0.2682	0.3609
FPA-SVM	Training Data	*MSE*	3.1259	2.6711	2.8385	4.4311
		SI	0.2118	0.2542	0.2189	0.3547
	Testing Data	*MSE*	3.4569	3.0356	3.3855	4.7001
		SI	0.2857	0.2641	0.2071	0.3801
PSO-SVM	Training Data	*MSE*	3.0383	2.2948	2.1935	4.1102
		SI	0.2569	0.2271	0.1738	0.3545
	Testing Data	*MSE*	3.0628	2.3825	3.2555	4.5208
		SI	0.2791	0.2249	0.2782	0.3712
GA-SVM	Training Data	*MSE*	3.1383	2.5949	2.1835	4.1309
		SI	0.2769	0.2378	0.1938	0.3645
	Testing Data	*MSE*	3.3428	2.4828	3.2051	4.7208
		SI	0.2762	0.2349	0.2782	0.3901
SVM-Defualt	Training Data	*MSE*	3.3259	2.8811	2.9335	4.5877
		SI	0.2758	0.2542	0.2589	0.4048
	Testing Data	*MSE*	3.8569	3.2357	3.4851	4.9127
		SI	0.2957	0.2674	0.2883	0.4069

TABLE 7. Summary the effects of training accuracy of different fault types.

Node ID	Data-loss Fault		Hardware Fault		Drift Fault		Gain Fault	
	IA	FPR	IA	FPR	IA	FPR	IA	FPR
001	89%	26%	99%	32%	87%	32%	92%	28%
002	88%	32%	98%	37%	84%	31%	90%	32%
004	90%	24%	89%	34%	85%	34%	92%	34%
030	93%	31%	93%	34%	86%	31%	93%	31%
031	87%	30%	97%	36%	85%	30%	87%	30%
051	88%	22%	96%	32%	86%	32%	91%	32%
052	89%	34%	89%	34%	85%	36%	89%	26%
078	92%	31%	89%	35%	97%	32%	88%	32%
099	91%	30%	88%	35%	89%	34%	85%	34%
AVG	**89.9%**	**28.9%**	**93.1%**	**34.3%**	**87.1%**	**32.4%**	**89.6%**	**31.0%**

known as the Hausdorff metric [25] is used to determine the distance difference between two datasets. Assumed, there two non-empty of subsets A and B; the distance between them calculated as follows.

$$D_H(A,B) = \max\{sup_{x\in A} inf_{y\in B} d(x,y), sup_{y\in B} inf_{x\in A} d(x,y)\} \quad (18)$$

where $D_H(A,B)$ is Hausdorff distance between two subset A and B; *sup* and *inf* are the supremum and infimum, respectively. FIGURE 9 and 10 depict the comparison of the proposed scheme for the classification of a collecting temperature dataset with the SVM, DT, and HMM methods on the faulty of Hardware errors and Data-loss Fault, respectively.

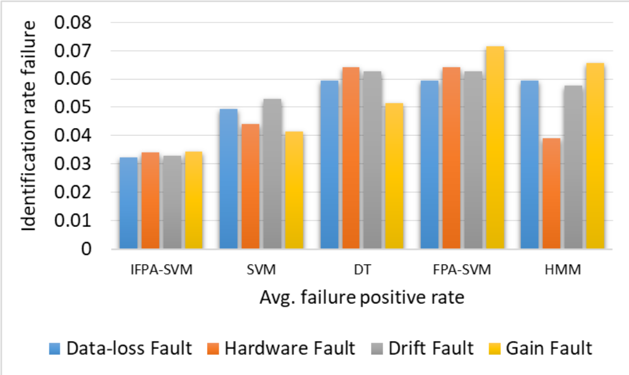


FIGURE 8. Comparison of the proposed IFPA-SVM for the classification of a collecting temperature dataset with the original FPA, SVM, DT, HMM methods concerning the FPR metric.

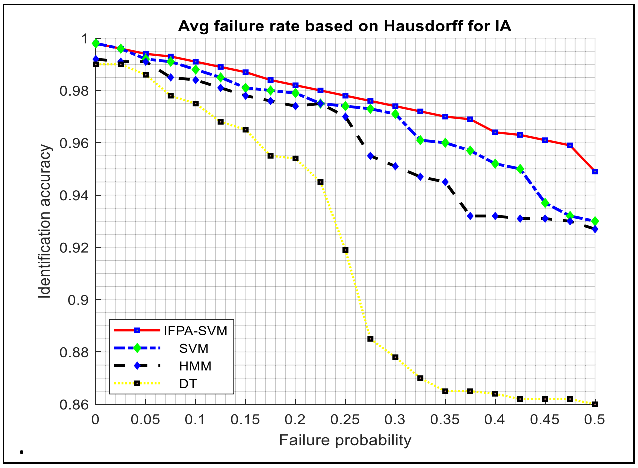


FIGURE 9. Comparison of the proposed method for the classification of collecting data with the SVM, DT, HMM on the hardware faulty.

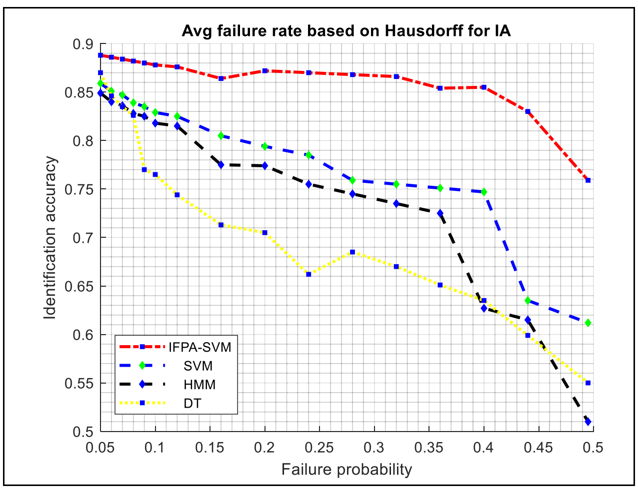


FIGURE 10. Comparison of the proposed method for the classification of collecting data with the SVM, DT, HMM on the data-loss fault.

It can be seen that our proposed method outperforms other competitors. The average of results the outcomes of the proposed method based on Hausdorff distance for IA is at 94.1 % and 89.0% on the hardware faulty and data-loss fault, respectively.

VI. CONCLUSION

In this paper, we presented a new scheme of collecting data classification for aggregating data in cluster heads (CHs) in WSN based on improving the classifier of the Support vector machine (SVM). We also proposed an improvement version for the Flower pollination algorithm (IFPA) to optimize parameters for the classification SVM. Due to the requirement of the precise data in several successful WSN applications, a decision function of classification should be deployed in CHs for identifying information correctly to aggregate the usual data for the next process.

The collecting environmental data like temperature, humidity, etc., are classified as collecting 'fault' or 'normal' data to aggregate and transfer them to the base station (BS). Compared with some existing methods, the proposed method offers an effective way of forwarding the correct data in WSN applications. The system design of the proposed scheme consists of majority components such as the collecting and pre-processing data, normalizing attributes of data, and training and testing datasets.

In the experimental section, the system was tested with collecting data by IFPA-SVM as improved classification. The experimental results were compared with the other methods in the literature of optimization algorithms and classifications, e.g., the original FPA, Genetic algorithm (GA), Particle swarm optimization (PSO), and labeling such as the support vector machine (SVM), decision tree (DT), hidden Markov model (HMM). The comparison result shows that the proposed method offers an effective way of forwarding the correct data for WSN applications. The system provides an accuracy of more than 97% throughout the data learning process. In data testing, the efficiency of the improved identification failure data offers more precise than the other competitors in comparison.

REFERENCES

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102–105, Nov. 2002.

[2] T.-T. Nguyen, J.-S. Pan, and T.-K. Dao, "An improved flower pollination algorithm for optimizing layouts of nodes in wireless sensor network," *IEEE Access*, vol. 7, pp. 75985–75998, 2019.

[3] T.-T. Nguyen, J.-S. Pan, and T.-K. Dao, "A compact bat algorithm for unequal clustering in wireless sensor networks," *Appl. Sci.*, vol. 9, no. 10, p. 1973, 2019.

[4] Z.-G. Du, J.-S. Pan, S.-C. Chu, H.-J. Luo, and P. Hu, "Quasi-affine transformation evolutionary algorithm with communication schemes for application of RSSI in wireless sensor networks," *IEEE Access*, vol. 8, pp. 8583–8594, 2020.

[5] D.-F. Ye, L.-L. Min, and W. Wang, "Design and implementation of wireless sensor network gateway based on environmental monitoring," in *Proc. Int. Conf. Environ. Sci. Inf. Appl. Technol.*, Jul. 2009, pp. 289–292.

[6] Y. Chen, S. Weng, W. Guo, and N. Xiong, "A game theory algorithm for intra-cluster data aggregation in a vehicular ad hoc network," *Sensors*, vol. 16, no. 2, p. 245, 2016.

[7] J.-S. Pan, P. Hu, and S.-C. Chu, "Novel parallel heterogeneous meta-heuristic and its communication strategies for the prediction of wind power," *Processes*, vol. 7, no. 11, p. 845, 2019.

[8] J.-S. Pan, T.-T. Nguyen, T.-K. Dao, T.-S. Pan, and S.-C. Chu, "Clustering formation in wireless sensor networks: A survey," *J. Netw. Intell.*, vol. 02, no. 04, pp. 287–309, 2017.

[9] T.-T. Nguyen, J.-S. Pan, and T.-K. Dao, "A novel improved bat algorithm based on hybrid parallel and compact for balancing an energy consumption problem," *Information*, vol. 10, no. 6, p. 194, 2019.
[10] S. G. Shiva Prasad Yadav and A. Chitra, "Wireless sensor networks—Architectures, protocols, simulators, and applications?: A survey," *Int. J. Electron. Comput. Sci. Eng.*, vol. 1, no. 4, pp. 1941–1953, 2012.
[11] T.-T. Nguyen, T.-K. Dao, M.-F. Horng, and C.-S. Shieh, "An energy-based cluster head selection algorithm to support long-lifetime in wireless sensor networks," *J. Netw. Intell.*, vol. 1, no. 1, pp. 23–37, 2016.
[12] C. F. García-Hernández, P. H. Ibarguengoytia-González, J. García-Hernández, and J. A. Pérez-díaz, "Wireless sensor networks and applications?: A survey," *J. Comput. Sci.*, vol. 7, no. 3, pp. 264–273, 2007.
[13] W. Guo, G. Chen, C. Yu, J. Su, and Z. Liu, "A two-stage clustering sleep scheduling algorithm with particle swarm optimization in wireless sensor networks," *Adhoc Sens. Wireless Netw.*, vol. 27, pp. 27–49, Jul. 2015.
[14] H. Cheng, Z. Xie, Y. Shi, and N. Xiong, "Multi-step data prediction in wireless sensor networks based on one-dimensional CNN and bidirectional LSTM," *IEEE Access*, vol. 7, pp. 117883–117896, 2019.
[15] S. Jia, L. Ma, and D. Qin, "Fault detection modelling and analysis in a wireless sensor network," *J. Sensors*, vol. 2018, Oct. 2018. Art. no. 7935802.
[16] F. Yuan, Y. Zhan, and Y. Wang, "Data density correlation degree clustering method for data aggregation in WSN," *IEEE Sensors J.*, vol. 14, no. 4, pp. 1089–1098, Apr. 2014.
[17] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[18] X.-S. Yang, "Flower pollination algorithm for global optimization," in *Proc. Int. Conf. Unconventional Comput. Natural Comput.*, vol. 7445, 2012, pp. 240–249.
[19] J.-S. Pan, L. Kong, T.-W. Sung, P.-W. Tsai, and V. Snasel, "Alpha-fraction first strategy for hierarchical wireless sensor networks," *J. Internet Technol.*, vol. 19, no. 6, pp. 1717–1726, 2018.
[20] H. C. Shih, J. H. Ho, B. Y. Liao, and J. S. Pan, "Fault node recovery algorithm for a wireless sensor network," *IEEE Sensors J.*, vol. 13, no. 7, pp. 2683–2689, Apr. 2013.
[21] C.-I. Wu, H.-Y. Kung, C.-H. Chen, and L.-C. Kuo, "An intelligent slope disaster prediction and monitoring system based on WSN and ANP," *Expert Syst. Appl.*, vol. 41, no. 10, pp. 4554–4562, Aug. 2014.
[22] A. Mahapatro and P. M. Khilar, "Fault diagnosis in wireless sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2000–2026, Apr. 2013.
[23] Z. Zhang, A. Mehmood, L. Shu, Z. Huo, Y. Zhang, and M. Mukherjee, "A survey on fault diagnosis in wireless sensor networks," *IEEE Access*, vol. 6, pp. 11349–11364, 2018.
[24] S. Zidi, T. Moulahi, and B. Alaya, "Fault detection in wireless sensor networks through SVM classifier," *IEEE Sensors J.*, vol. 18, no. 1, pp. 340–347, Jan. 2018.
[25] K. Flouri, B. Beferull-Lozano, and P. Tsakalides, "Training a SVM-based classifier in distributed sensor networks," in *Proc. 14th Eur. Signal Process. Conf.*, Sep. 2006, pp. 1–5.
[26] A. Y. Barnawi and I. M. Keshta, "Energy management in wireless sensor networks based on naive bayes, MLP, and SVM classifications: A comparative study," *J. Sensors*, vol. 2016, Feb. 2016, Art. no. 6250319.
[27] H. Zhang and Z. Pan, "Cross-voting SVM method for multiple vehicle classification in wireless sensor networks," *Sensors*, vol. 18, no. 9, p. 3108, 2018.
[28] S. Fine, Y. Singer, and N. Tishby, "The hierarchical hidden Markov model: Analysis and applications," *Mach. Learn.*, vol. 32, no. 1, pp. 41–62, 1998.
[29] J. He, H. Li, and J. Tan, "Real-time daily activity classification with wireless sensor networks using hidden Markov model," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 3192–3195.
[30] P. Tang and T. W. S. Chow, "Wireless sensor-networks conditions monitoring and fault diagnosis using neighborhood hidden conditional random field," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 933–940, Jun. 2016.
[31] X. Song, G. Chen, and X. Li, "A weak hidden Markov model based intrusion detection method for wireless sensor networks," in *Proc. Int. Conf. Intell. Comput. Integr. Syst.*, Oct. 2010, pp. 887–889.
[32] C. Yang, C. Liu, X. Zhang, S. Nepal, and J. Chen, "A time efficient approach for detecting errors in big sensor data on cloud," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 2, pp. 329–339, Feb. 2015.
[33] A. Garofalo, C. Di Sarno, and V. Formicola, "Enhancing intrusion detection in wireless sensor networks through decision trees," in *Proc. Eur. Workshop Dependable Comput.*, 2013, pp. 1–15.
[34] B. Sugiarto and R. Sustika, "Data classification for air quality on wireless sensor network monitoring system using decision tree algorithm," in *Proc. 2nd Int. Conf. Sci. Technol.-Comput. (ICST)*, Oct. 2016, pp. 172–176.
[35] J. W. Barron, A. I. Moustapha, and R. R. Selmic, "Real-time implementation of fault detection in wireless sensor networks using neural networks," in *Proc. 5th Int. Conf. Inf. Technol., New Generat. (ITNG)*, Apr. 2008, pp. 378–383.
[36] P. Ballal, A. Ramani, M. Middleton, C. Mcmurrough, A. Athamneh, W. Lee, C. Kwan, and F. Lewis, "Mechanical fault diagnosis using wireless sensor networks and a two-stage neural network classifier," in *Proc. IEEE Aerosp. Conf.*, Mar. 2009, pp. 1–10.
[37] T. Muhammed and R. A. Shaikh, "An analysis of fault detection strategies in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 78, pp. 267–287, Jan. 2017.
[38] R. Sathiyavathi and B. Bharathi, "A review on fault detection in wireless sensor networks," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2017, pp. 1487–1490.
[39] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.
[40] X.-F. Yan, J. Yu, and F. Qian, "Adaptive mutation differential evolution algorithm and its application to estimate soft sensor parameters," *Kongzhi Lilun Yu Yingyong/Control Theory Appl.*, vol. 23, no. 5, pp. 744–748, 2006.
[41] J. J. Liang, B. Y. Qu, and P. N. Suganthan, "Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization," in *Proc. Comput. Intell. Lab. Zhengzhou Univ. Zhengzhou China Tech. Nanyang Technol. Univ.*, Singapore, 2013, p. 635.
[42] G.-R. Xue, C. Lin, Q. Yang, W. Xi, H.-J. Zeng, Y. Yu, and Z. Chen, "Scalable collaborative filtering using cluster-based smoothing," in *Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, 2005, pp. 114–121.
[43] A. Stoffelen, "Toward the true near-surface wind speed: Error modeling and calibration using triple collocation," *J. Geophys. Res., Oceans*, vol. 103, no. 4, pp. 7755–7766, Apr. 1998.
[44] F. Kuang, W. Xu, and S. Zhang, "A novel hybrid KPCA and SVM with GA model for intrusion detection," *Appl. Soft Comput.*, vol. 18, pp. 178–184, May 2014.
[45] C.-L. Huang and J.-F. Dun, "A distributed PSO–SVM hybrid system with feature selection and parameter optimization," *Appl. Soft Comput.*, vol. 8, no. 4, pp. 1381–1391, Sep. 2008.
[46] W. Li, D. Kong, and J. Wu, "A new hybrid model FPA-SVM considering cointegration for particular matter concentration forecasting: A case study of kunming and yuxi, China," *Comput. Intell. Neurosci.*, vol. 2017, Aug. 2017, Art. no. 2843651.

THI-KIEN DAO received the Ph.D. degree in electronic and engineering from the National Kaohsiung University of Technology and Sciences, Taiwan, in 2019. She is currently a Lecturer with the College of Information Science and Engineering, Fujian University of Technology. Her current research interests include computational intelligence, data mining, and sensor networks.

TRONG-THE NGUYEN received the Ph.D. degree in communication engineering from the National Kaohsiung University of Applied Sciences, Taiwan, in 2016. He is currently a Lecturer with the College of Information Science and Engineering, Fujian University of Technology, and with the Department of Information technology, Haiphong University of Manager and Technology, Vietnam. His current research interests include computational intelligence, signal processing, and sensor networks.

JENG-SHYANG PAN received the B.S. degree in electronic engineering from the National Taiwan University of Science and Technology, in 1986, the M.S. degree in communication engineering from National Chiao Tung University, Taiwan, in 1988, and the Ph.D. degree in electrical engineering from the University of Edinburgh, U.K, in 1996. He is currently a Professor with the College of Computer Science and Engineering, Shandong University of Science and Technology. He joined the Editorial Board of *LNCS Transactions on Data Hiding and Multimedia Security*, the *Journal of Computers*, and the *Chinese Journal of Electronics*. His current research interests include soft computing, information security, and signal processing.

YU QIAO received the B.S. degree in electronic science and technology from the Fujian University of Technology, China, in 2018. He is currently working with the Shandong University of Science and Technology, Qingdao, China. His current research interests include swarm intelligence, computer vision, and data mining.

QUOC-ANH LAI received the bachelor's degree from the Faculty of Information Technology, University of Engineering and Technology, Vietnam National University, Hanoi, in 1998, and the master's degree from the Computer Science Department, Posts and Telecommunications Institute of Technology Hanoi, in 2013. He is currently a Researcher with the Department of Pattern Recognition and Image Processing, Institute of Information Technology, Vietnam Academy of Science and Technology, Vietnam. His research interests include pattern recognition and image processing, computational intelligence, and natural language processing.