Arab J Sci Eng (2017) 42:3435–3449
DOI 10.1007/s13369-017-2524-0

RESEARCH ARTICLE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

A Novel LWCSO-PKM-Based Feature Optimization and Classification of Attack Types in SCADA Network

Dhanalakshmi Krishnan Sadhasivan[1] · **Kannapiran Balasubramanian**[2]

Received: 25 June 2016 / Accepted: 10 April 2017 / Published online: 24 April 2017

Abstract Currently, Supervisory Control and Data Acquisition (SCADA) systems are widely used in the remote monitoring and control of the large-scale manufacturing plants and power grids. The development of high-security SCADA is the major requirement due to their vulnerability to attacks based on the architectural constraints. The decision making regarding the controlling of power flows and the replacement of faulty devices is based on the two stages normal or attacked. The observations from the sensor play the major role in the classification of normal and abnormal patterns. With the increase in a number of observations, the dimensionality of features is high and thus there is a chance of misleading results during the classification progress. Various classification and the intrusion detection (ID) algorithms are available to reduce the dimensionality of features for better classification. This paper proposes a novel approach for feature optimization and classification of the attack types in the SCADA network with better performance than the existing algorithms. The Linear Weighted Cuckoo Search Optimization (LWCSO) algorithm in proposed work selects the best features from the overall features. A Probabilistic Kernel Model (PKM) updates the weight function of each node to form the clusters representing the optimal features. The label is applied to each cluster based on the difference between the set of labeled training features with the testing feature set. Based on this label, the features are applied to detect the anomaly node in the network area. From the classification result, if the attack type is already known, then appropriate action is taken immediately. If the attack type is unknown, its type is added to the database. The periodical discovery of the type of attack and the database update with the unknown attacks increases the detection ability effectively. From the performance analysis, it is observed that the proposed LWCSO-PKM approach achieves better performance than the existing classification techniques and IDS algorithms.

Keywords Intrusion detection system · Linear weighted CSK optimization (LWCSO)Algorithm · Probabilistic Kernel Model (PKM) · Supervisory Control And Data Acquisition (SCADA)System

1 Introduction

Numerous advancements in the computer vision-based monitoring and controlling cause the number of hacking and intrusion attacks and hence the security assurance is the major need of networking system [1]. The prediction of abnormal data flow and the isolation of normal from the abnormal with the reduced features plays the major role in decision making regarding the security assurance. Currently, many researchers focus on the development of Intrusion Detection System (IDS) to predict and classify the normal and abnormal nodes. The main objective of the IDS is to protect the confidentiality and integrity of the network from the risk of attacks by the intruders or hackers and ensure the availability of data [2]. The IDS is classified into two categories such as anomaly detection and signature-based detection systems [3]. The continuous learning and less maintenance are the merits of anomaly-based IDS. However, the anomaly-based

✉ Dhanalakshmi Krishnan Sadhasivan
k.s.dhanalakshmi@klu.ac.in

Kannapiran Balasubramanian
b.kannapiran@klu.ac.in

[1] Department of ECE, Kalasalingam University, Krishnakoil, Tamil Nadu 626126, India

[2] Department of Instrumentation and Control Engineering, Kalasalingam University, Krishnakoil, Tamil Nadu, India

IDS never send an alarm even if the malicious node looks like a normal node that leads to misclassification. The signature-based IDS detect the attacks based on the signature of the previous attacks which is considered as a more accurate system for attack detection. However, the signature-based IDS can detect the attack only if the patterns are matched with the database. The increase in a number of signatures limits the efficiency of the signature-based IDS adversely.

1.1 Motivation

Industrial network security is currently emerging field with the rapid expansion of computer networks all over the world. Improvement in the security of the industrial network involves detection and prevention of the intrusion attacks on the computers in the network. The main objective of the industrial network security is to protect the SCADA system. SCADA is a technology used in the industrial network for collecting data from the remote facilities and send control instructions to the facilities. It enables a processing center to automatically control a large-scale industrial process such as oil pipeline system, manufacturing plants, refineries or power grid. Increase in the security ensures reliable operation of the SCADA system [4]. Anomaly and signature-based intrusion detection techniques are the main techniques that are used for the detection of the SCADA-specific attacks [5]. Almalawi et al. [6] proposed an anomaly-based intrusion detection approach to detect the cyber-attacks in the SCADA system. The signature-based IDS depend on the matching observations to the abnormal patterns. Snort [7] is a signature-based IDS used to detect SCADA-specific attacks with predefined rules. Yang et al. [8] employed the Snort IDS and proposed signature-based rules to detect the attacks. But, the rule-based IDS suffers from computational complexity and the presence of huge size-features limited the classification performance. Most of the signature-based IDSs respond passively to the attacks and do not mitigate the malicious impacts. Hence, design of an efficient IDS to counter the attacks in the SCADA system is an important topic of interest for the researchers.

1.2 Our Proposed Work

To overcome the dimensionality limitations in the existing IDSs, our research work proposes a novel optimal feature optimization and classification model of the attack types in the SCADA network. The LWCSO algorithm selects the best features from the overall feature set that correspond to the name of the attack from the table list. A novel method of Kernel function is updates the weight function of each node and form clusters of optimal feature data. Also, the difference between the attack information extracted with the table information applies the label to each cluster. Based on this label, the features are applied to detect anomaly node in the network area. A novel PKM-based classifier used in proposed work classifies the packets arrived from the particular node as either normal or attack. If the packet flow is detected as a new type of attacker, its label is updated in the library. The attack level is verified, and feature matrix and label are updated. Our proposed LWCSO-PKM algorithm achieves better intrusion detection results by using the label information when compared to traditional IDS. The type of attack estimation and verification from a predefined dataset by using the classification algorithm is the major contribution of proposed work. The type of attack is estimated from a predefined dataset and type of the attack is verified by using the classification algorithm. From the classification result, if the attack type is already known, then appropriate action is taken to the predicted node. But, if the attack type is unknown, its type is updated to the database and a new label is assigned to it. The performance analysis shows that the proposed LWCSO-PKM approach yields higher accuracy, sensitivity, specificity, precision, recall, Jaccard Coefficient, Dice coefficient, Kappa Coefficient, detection rate and lower false alarm rate than the existing classification techniques and IDS algorithms.

1.3 Organization

The organization of the remaining sections of the paper is described as follows: Sect. 2 presents a brief review of the existing research articles related to the IDS for SCADA systems. Section 3 explains the proposed work in detail including the LWCSO algorithm and PKM-based classification algorithm. Section 4 illustrates the comparative analysis of the proposed LWCSO-PKM-based feature selection and classification approach with the existing Support Vector Machine (SVM) classifier. Section 5 gives an overview of the conclusion and future implementation of the proposed work.

2 Related Works

This section gives a brief review of the existing research works related to the IDS for SCADA systems. Yang et al. [4] presented a novel approach for a next-generation SCADA-specific IDS. The proposed system analyzed multiple attributes to provide a complete solution for mitigating the cyber-attack threats. A multilayer cyber-security structure was proposed for the protection of the smart grids without compromising the availability of normal data. Maglaras and Jiang [9] presented an intrusion detection module based on the combination of One-Class Support Vector Machine (OCSVM) with Radial Basis Function (RBF) Kernel and recursive k-means clustering. The OCSVM module was trained by the network traces offline to detect anomalies in the

real-time SCADA system. Fahad et al. [10] proposed a novel approach for the efficient and accurate identification of the best features by combining the results of some feature selection techniques to find the consistent features, and use the proposed concept of support to select a smallest set of features and cover data optimality. The proposed approach achieved high accuracy and run-time performance of the classifier. Gong et al. [11] proposed a feature selection approach to improve the detection reliability of the Multi-Agent Intrusion Detection System (MIDS) architecture. MIDS was designed to ensure decentralized intrusion detection and protect the Industrial Control System (ICS). The proposed approach achieved high true positive rate and low false positive rate when compared with other feature selection algorithms.

Nader et al. [12] investigated the use of one-class classification algorithms such as Support Vector Data Description (SVDD) and Kernel Principle Component Analysis (PCA) for intrusion detection in SCADA systems. The optimal choice of the bandwidth parameter in the RBF kernels was found out. Erez and Wool [13] described a novel domain-aware anomaly detection system that detects irregular changes in Modbus/Transfer Control Protocol (TCP) SCADA control register values. An automatic classifier was developed to identify the classes of registers such as sensor registers, counter registers and constant registers. The classifier achieved high true positive rate of 93% and low false alarm rate of 0.86% for the correctly classified registers. Moon et al. [14] proposed an IDS based on the decision tree through the analysis of behavior information to detect Advanced Persistent Threats (APT) attacks. The possibility of the initial intrusion was detected and damage to the network was reduced by quickly responding to the APT attacks.

Lin et al. [15] introduced a semantic analysis framework by combining the knowledge of both cyber and physical infrastructure based on the distributed network of (IDSs). The proposed framework provided reliable detection of the malicious commands within a small time. Mo et al. [16] developed model-based techniques for detecting the integrity attacks on the sensors of the control system. The probability of attack detection was optimized by improving the control performance. Gao and Morris [16] described a set of individual and state-based IDS rules to detect the cyber-attacks and store the proof of the attacks for the future analysis. The detection rate of the IDS rules was improved. Wang et al. [17] proposed a new relation-graph-based scheme for detecting false data injection attacks in the SCADA system. A novel detection model was designed to conclude the attack origin in the SCADA system. The proposed scheme achieved low false positive rate and accurate detection of the attack origins.

Rusu et al. [18] introduced a novel methodology to automatically configure Snort-based Anomaly Detection Systems (ADSs) deployed in the SCADA systems. The proposed methodology achieved better performance with large network topologies. Jin et al. [19] presented a Bayesian anomaly IDS using fuzzy probability assignment. The real-time probabilities of security-breaching events were calculated to decide whether the SCADA system is under attack. The probabilities of the security events were described accurately by using the continuous fuzzy probability model. Nasr and Varjani [20] proposed a novel alarm-based statistical detection method for identifying potential insider attacks at the substations and total transmission system in the power grid. The anomalies were detected by using a minimum number of alarms. McLaughlin et al. [21] presented an integrated approach for detecting the cyber-attacks to the modern ICS.

Jiang and Yasakethu [22] applied the concept of one-class SVM to detect the unusual patterns from the inputs and generate alarms for the on-site engineers for further investigation. The proposed algorithm achieved high detection rate and excellent protection of the SCADA systems. Do et al. [23] proposed a Finite Moving Average (FMA) algorithm to detect the covert attack on a water distribution system. The proposed algorithm achieved reduction in the worst-case probability of false alarm. Hug and Giampapa [24] introduced analytical techniques to perform vulnerability analysis on the SCADA system of the power grid. Yang et al. [25] presented a cyber-security test bed for investigating an Address Resolution Protocol (ARP) spoofing based man-in-the-middle attack in the SCADA systems. The main drawbacks of the existing SCADA-based IDS are high false positive rate. Also, the periodical update of attack library is the major need to detect unknown attacks exactly. The size of the feature set is a rough indicator of the efficiency for the feature selection-based IDS approaches. Large size feature sets require more memory and iteration time for classification. More features do not necessarily give better results. Hence, feature selection is a key research problem in the IDS. Also, there is a need to improve the capabilities of the proposed IDS to detect and prevent the emerging threats. To overcome these drawbacks, this paper proposes an LWCSO-PKM-based feature selection and classification of the attack types.

3 Proposed LWCOS-PKM-Based Feature Selection and Classification of Attack Types

This section explains the proposed work including the LWCSO algorithm and PKM-based classification algorithm. Initially, preprocessing of the input dataset is performed to arrange the attributes in the training dataset. A novel LWCSO algorithm selects the optimal features according to the fitness value of the feature attributes. A PKM-based classification uses a novel model of Kernel function to estimate the matching point between training feature set and the extracted feature vector. Based on the matched result, the proposed work assigns the label to the feature vectors. Based on the

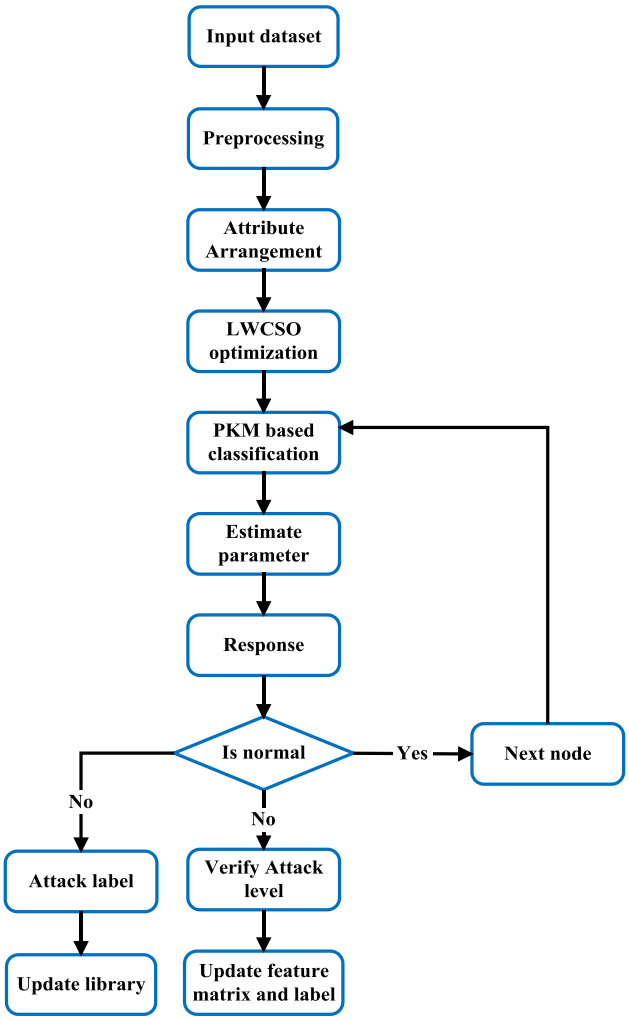

Fig. 1 Flow diagram of the overall LWCSO-PKM-based feature optimization and classification of attack types in SCADA

status of classification of packet flows among the nodes, three processes are performed. If the packet flow is normal, then the next state packet flow is analyzed. The level of attack is verified if there is any abnormality in the packet flow. If there is any unknown attack, then the library is updated with their information for the future immediate decisions. Figure 1 shows the flow diagram of the overall LWCSO-PKM-based feature optimization and classification of attack types in the SCADA. In our proposed optimization technique, we present a new objective function to select the relevant features that estimates the variation in the feature set and forms a cluster to find the best fitness value. Also, the selection of best attributes is based on the response of each sensor connected with the control unit. According to that variation, the best fitness is extracted and the indices corresponding to the fitness value are selected. The selection process is based on the criteria of selected indices are higher than the average value of selected indices. In the PKM classifier method, the features are trained by using a new Kernel model that estimates the difference in each class of the given feature set, forms it as a cluster and labels the cluster. After clustering, the rules are formed for the classification of testing feature set. The labeled trained feature utilization in PKM classifier provided better performance for the given testing feature set than the traditional classification algorithms.

This paper proposes a new optimization and classification model for the SCADA system. The data are obtained from four relays. Totally, 124 features are obtained for each data. The direct extraction of all the features increased the memory space and time consumption. To overcome this issue, the proposed work selects the optimal feature attributes from the overall attributes using an enhanced Cuckoo Search Optimization (CSO) called as LWCSO for feature selection. For the selection of optimal features, we estimate the repeated contents and attributes containing null/same values for abnormalities prediction. In this case, it is difficult to separate the normal and malicious nodes. The CS optimization selects the best features based on the fitness value by using our objective function. The objective function is determined based on the relay signals and cost function. If the cost function is low, we can easily analyze the clustering and classification of the normal and attack attributes.

This paper enhances the optimization function based on the objective function. When compared with the traditional CSO, our enhanced LWCSO algorithm forms the objective function based on the relay signals and response from the sensors in power grids. The sensors in the SCADA system obtain data for each and every time sample. Here, we extract the number of observations for more number of time samples. The loading of complete dataset in training process maximizes the memory and processing time for the classification. The proposed LWCSO algorithm reduces both the memory and time consumption by selecting the optimal feature set. Among the 124 features, 50 features are selected as the best features through the proposed LWCSO algorithm. Hence, the performance of PKM classifier is better due to the reduction in the total number of features. The classification model is enhanced by implementing a new Kernel model to estimate and learn the best range. Our proposed Kernel ensures better learning than the existing Gaussian and Bayesian Kernels. The kernel formulation forms the rule to determine whether the node is a normal or anomalous. Finally, our proposed PKM classification provides better performance based on the optimal features than the existing classification algorithms.

3.1 LWCSO

Cuckoo Search (CS) [26,27] is a population-based optimization algorithm that is formulated based on the breeding behavior of cuckoo. The general rules in the CS approach are

- Each cuckoo lays only one egg at a time, and dumps its egg in a randomly chosen nest.
- The nests with high quality of eggs will proceed over to the next generations.
- The number of available host nests is fixed, and the egg laid by a cuckoo is discovered by the host bird with a probability $p_a \in [0, 1]$. In this case, the host bird can either throw the egg away or abandon the nest, and build a completely new nest.

Each egg in a nest represents a new optimal solution. The number of parameters to be tuned in the CS algorithm is lesser than the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). Hence, CS is adapted to a wider range of optimization problems.

The LWCSO algorithm presents an optimization technique for feature selection. For this process, a novel model of objective function is modified in the cuckoo search optimization. This objective function extracts weight updating for each iteration. This verifies whether the particles lie in the maximum fitness value or not. If the profit is maximum than the previous value, the fitness value is updated. Otherwise, the previous value is maintained. According to the profit, the fitness value for each iteration is extracted and coordinates of the particles are estimated. If it is greater than centroid, then the feature attribute is selected. Otherwise, the attribute is not selected. Thus, the optimization performs feature selection according to the fitness value of feature attributes.

The random center position of the cuckoo is extracted based on the maximum and minimum values of the training features f(x) by using the following equation

$$C_{\text{center}} = ((\max(f(x)) - \min(f(x)) * \text{Rand}) + \min(f(x))) \quad (1)$$

where 'Rand' denotes the random value ranging from 0 to 1. The initial fitness value is estimated based on the upper and lower limits of the training features and random particles.

$$\text{Best}_{\text{fit}} = \left(\left(\text{Var}_{\text{high}} - \text{Var}_{\text{low}}\right) * f(y)_{1,2,\ldots\text{npar}}\right) + \text{Var}_{\text{low}} \quad (2)$$

where $f(y)$ denotes the random particles, Var_{high} is the upper limit of $f(x)$, Var_{low} is the lower limit of $f(x)$ and npar is the number of feature particles. The cuckoo particles are based on the feature attributes.

The maximum profit of the initial iteration is estimated by using the objective function defined as

$$\text{Max}_{\text{Pro}} = W + \sum_{i=1}^{N} f(x)_i^2 + \left(\text{Sin}\left(2 * pi * f(x)_i\right)\right) \quad (3)$$

where 'W' denotes the maximum weight of the cuckoos. The cluster formation is defined by updating the objective function. The number of eggs laid in the allocated area is estimated by using the following equation

$$\text{Cuckoo}_{\text{Egg}} = \left(\left(\text{Egg}_{\text{Max}} - \text{Egg}_{\text{Min}}\right) * \text{Rand}\right) + \text{Egg}_{\text{in}} \quad (4)$$

The 'X' and 'Y' coordinates describing the location of cuckoo for $n = 1, 2\ldots$ number of iteration are given by

$$X_{\text{Co-ordinate}}(n) = x(n-1) + \left(\left(\text{Rand}^{-\frac{1}{alpha}}\right) * \cos(\text{Rand} * 2 * pi)\right) \quad (5)$$

$$Y_{\text{Co-ordinate}}(n) = y(n-1) + \left(\left(\text{Rand}^{-\frac{1}{alpha}}\right) * \cos(\text{Rand} * 2 * pi)\right) \quad (6)$$

The radius of egg laying region of the cuckoo is computed based on the probability $P(m)$ of number of eggs at each center location and upper and lower limits of the training features.

$$\text{Cuckoo}_{\text{Radius}} = P(m) * \left(\text{Radius} * \left(\text{Var}_{\text{high}} - \text{Var}_{\text{low}}\right)\right) \quad (7)$$

The radius of the egg laying region is updated as

$$\text{Radius}_{\text{Update}}(l) = (-1)^{\text{Rand}_{0,1}} * \text{Radius}_{\text{Pre}} * \cos(\text{Radius}_{\text{Num}} * \left(\frac{2}{\text{Radius}_{\text{Num}} - 1}\right) + \text{Radius}_{\text{Pre}} * \sin(\text{Radius}_{\text{Num}} * \left(\frac{2}{\text{Radius}_{\text{Num}} - 1}\right) \quad (8)$$

The maximum profit from the objective function is updated

$$\text{Max}_{\text{Pro\_update}} = W + \sum_{i=1}^{N} C(x)_i^2 + \left(\text{Sin}\left(2 * pi * C(x)_i\right)\right) \quad (9)$$

where $C(x) = \left\{\text{Cuckoo}_{\text{Population}}, \text{Egg}_{\text{Position}}\right\}$

$C(x)$ is the set of cuckoo population and egg position. If the maximum updated profit is greater than the maximum profit, the maximum updated profit is updated and the center position of the cuckoo is updated to a new position. Otherwise, the same center position is maintained. The updated profit value and cuckoo center is used for the formation of clusters, until the predefined number of iterations is reached. The updated cuckoo center is estimated as the best fitness value output 'BF'. The selected features are obtained, when the best fitness value is greater than the average level of the best fitness value.

LWCSO Optimization algorithm

Input: Feature matrix 'T'
Output: Selected Feature, 'ST'
Step 1: Initialize cuckoo as providing training features, $f(x)$ and Accuracy value
Step 2: Extract random center position of cuckoo by using Eq. (1)
Step 3: Extract Initial fitness value using Eq. (2)
Step 4: Estimate maximum profit of initial iteration using Eq. (3)
Step 5: Extract cluster formation with objective function updation
For i = 1 to Number of iteration
If ($i \leq \text{Max}_{\text{iter}}$) && ($\text{Max}_{\text{Pro}} >$ Accuracy)
Estimate number of eggs laying in allocated area by using Eq. (4)
Estimate 'X' and 'Y' coordinates of cuckoo by using Eqs. (5) and (6)
Calculate radius of egg laying region of cuckoo by using Eq. (7)
Update Radius of egg laying region by using Eq. (8)
Update maximum profit from objective function by using Eq. (9)
If ($\text{Max}_{\text{Pro\_update}} > \text{Max}_{\text{Pro}}$)
Update $\text{Max}_{\text{Pro\_update}}$;
Update $\text{CK}_{\text{center}}$ to New Position
End If
Step 6: Updated profit and cuckoo center is used for cluster formation until the iteration size is reached.
Step 7: Extract Updated Cuckoo Center as Best Fitness value output $'BF'$.
Step 8: $ST = T(R, BF > \text{avg}(BF))$

3.2 PKM-Based Classification

The PKM-based classification process uses a novel model of Kernel function to estimate the matching point of feature set from the training and extracted feature vector. The distance between the features and standard deviation of the features are estimated. Then, the probability to extract the matching level of attributes among the feature set is estimated. The maximum matching probability results in the generation of the classified output. This returns the label information according to the matching feature of the training set.

The selected feature is used as an input for the classification process. The probability array π is initialized as

$$\pi = P(q_1 = s_i) \tag{10}$$

where, 's' is the state of the training set for $i = 1, 2 \ldots N$, 'N' is the size of the training set 'ST' and 'q' is the fixed state sequence for the length of the testing feature 'V'. The attributes of the training set are extracted corresponding to the row and column size of the testing feature. The distance between the training and testing features is computed as

$$d_i = \sqrt{(s_i - V_i)^2 + \left(s_j - V_j\right)^2} \tag{11}$$

The corresponding labels of the training set are extracted as function of the computed distance value. The probability array is calculated based on the labels of the training set and length of the training set.

$$\pi_i = \frac{\sum_{k=1}^{m} s_i\left(q_{i,j}(d)\right)}{m} \tag{12}$$

The standard deviation of the feature matrix σ_T and testing features σ_V is calculated by using the following equations

$$\sigma_T = \sqrt{\frac{1}{m}\sum_{i=1}^{N}\left(s_i\left(q_{i,j}(d)\right) - \pi_i\right)^2} \tag{13}$$

$$\sigma_V = \sqrt{\frac{1}{n}\sum_{i=1}^{N}\left(V_i\left(q_{i,j}(d)\right) - \pi_i\right)^2} \tag{14}$$

where 'm' denotes the length of ''and 'n' denotes the length of 'V_i'. The probability for the training feature and testing feature is estimated with respect to the size (N) of the training set and size (M) of the testing feature.

$$P(STr|\pi_i) = \left(2\prod\sigma_T^2\right)^{\frac{-N}{2}} * \mathrm{e}^{\left\{\left(\frac{-1}{2\sigma_T^2}\right)\|STr-\pi_i\|^2\right\}} \tag{15}$$

$$P(V) = \left(2\prod\sigma_D^2\right)^{\frac{-M}{2}} * \mathrm{e}^{\left\{\left(\frac{-1}{2\sigma_D^2}\right)\|V_i-\pi_i\|^2\right\}} \tag{16}$$

If the probability of the training feature is greater than the probability of the testing feature, the classified label for the testing feature is obtained.

PKM classification algorithm

Input: Updated Training set, 'ST', Testing Feature, 'V', and Label, 'L' .
Output: Classified Label, 'CL' .
Step 1: Initialization of probability array
Step 2: For (i = 1 to Row_size (V))
Step 3: For (j = 1 to Column_size (V))
Step 4: $s_i = STr\,(i, j)$; //Extract attributes of training set
Step 5: Compute distance between attributes and features by using Eq. (11)
Step 6: $q_{i,j} = L\,(d_i)$; //Extract corresponding labels of training set
Step 7: Calculate probability array by using Eq. (12)
Step 8: Calculate standard deviation of the feature matrix and testing features by using eqns (13) and (14)
Step 9: Estimate probability for training feature using Eq. (15)
Step 10: Estimate probability for testing feature using Eq. (16)
Step 11: ***If*** $(P\,(STr|\pi_i) > P\,(V))$ //Condition for feature verification
Step 12: $CL_i = L\,(P\,(V))$; //Classified label
Step 13: ***End if***
Step 14: End 'j' loop
Step 15: End 'i' loop

4 Performance Analysis

This section illustrates the comparative analysis of the proposed LWCSO-PKM-based feature selection and classification approach with the existing SVM classifier [9]. The PKM-based classification process uses a novel model of Kernel function to estimate the matching point of feature set from the training and extracted feature vector. The distance between the features and standard deviation of the features are estimated. Then, the probability to extract the matching level of attributes among the feature set is estimated. The maximum matching probability results in the generation of the classified output. This returns the label information according to the matching feature of the training set. If the node is detected as a new type of attacker, its label is updated in the library. The attack level is verified and feature matrix and label are updated. In our proposed work, both feature optimization algorithm and classifier achieves better intrusion detection results with the label information when compared to traditional IDS. The type of attack is estimated from a predefined dataset and type of the attack is verified by using the classification algorithm.

4.1 Dataset Description

The Mississippi State University (MSU) SCADA security laboratory database [28,29] is used for evaluating the performance of the proposed LWCSO and PKM algorithms. The security laboratories are built with the commercial hardware and software devices from multiple vendors. The test bed includes commercial equipment and software to monitor and control the laboratory-scale physical processes from multiple critical infrastructures. The test bed is used for the pedagogical and research purposes. The 37 scenarios are categorized as eight natural events, twenty-eight attack events and one null event. There are 41 types of attack events. Table 1 depicts the three-class classification of the attack events.

The Australian Defense Force Academy-Linux Dataset (ADFA-LD) [30,31] is used for the performance evaluation. This dataset uses a fully patched Ubuntu Linux 11.04 [32] as the host operating system. This dataset provides an existing Linux dataset for evaluating the performance of the proposed work with respect to the traditional Host-based IDS (HIDS). It provides a better indication of performance against the attacks. The ADFA dataset includes totally 1000 observations. Among them, 757 observations are found as attack and 243 observations are normal. The types of the attacks are analysis, backdoor, Denial of Service (DoS), exploits, Fuzzers, reconnaissance, shellcode and worms. The MSU SCADA security laboratory dataset includes 5069 observations. Out of them, 1544 observations are found as natural and 3525 observations are found as attack. Table 2 shows the number of observations in dataset. Table 3 shows the attack events in the ADFA-LD dataset. Tables 4, 5 and 6 describe the natural events, no event and attack event scenarios in the MSU SCADA security laboratory dataset.

Table 1 Three-class classification of attack events

	Attack events	Natural events	No events
Scenarios	7, 8, 9, 10, 11, 12, 15, 1, 6, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 36, 37, 38, 3, 9, 40	1, 2, 3, 4, 5, 6, 13, 14	41

Table 2 Number of observations in dataset

	Mississippi power bus dataset	ADFA Linux Dataset
Total number of observations	5069	1000
Number of attack observations	3525	757
Number of normal observations	1544	243

Table 3 Attack events in ADFA-LD dataset

Attack type	Number of attack events
Analysis	14
Exploits	332
Reconnaissance	114
Shellcode	15
Worms	2
DoS	124
Backdoor	6
Fuzzers	150
Total attack events	757

Table 4 Natural event scenarios in MSU SCADA security laboratory dataset

Scenario	Natural events (SLG faults)
Natural events	
1	Fault from 10–19% on Line 1 (L1)
2	Fault from 20–79% on L1
3	Fault from 80–90% on L1
4	Fault from 10–19% on L2
5	Fault from 20–79% on L2
6	Fault from 80–90% on L2
Natural events (line maintenance)	
13	Line L1 maintenance
14	Line L2 maintenance

Table 5 No event scenario

Regular operation Scenario	No events (normal operation)
41	Normal operation load changes

Table 6 Attack event scenarios

Scenario	Attack type
Data injection	
Attack sub-type (SLG fault replay)	
7	Fault from 10–19% on L1 with tripping command
8	Fault from 20–79% on L1 with tripping command
9	Fault from 80–90% on L1 with tripping command
10	Fault from 10–19% on L2 with tripping command
11	Fault from 20–79% on L2 with tripping command
12	Fault from 80–90% on L2 with tripping command
Remote tripping command injection	
Attack sub-type (Command injection against single relay)	
15	Command Injection to R1
16	Command Injection to R2
17	Command Injection to R3
18	Command Injection to R4
Attack Sub-type (Command injection against single relay)	
19	Command Injection to R1 and R2
20	Command Injection to R3 and R4
Relay setting change	
Attack Sub-type (disabling relay function—single relay disabled & fault)	
21	Fault from 10–19% on L1 with R1 disabled & fault
22	Fault from 20–90% on L1 with R1 disabled & fault
23	Fault from 10–49% on L1 with R2 disabled & fault
24	Fault from 50–79% on L1 with R2 disabled & fault
25	Fault from 80–90% on L1 with R2 disabled & fault
26	Fault from 10–19% on L2 with R3 disabled & fault
27	Fault from 20–49% on L2 with R3 disabled & fault
28	Fault from 50–90% on L2 with R3 disabled & fault
29	Fault from 10–79% on L2 with R4 disabled & fault
30	Fault from 80–90% on L2 with R4 disabled & fault
Attack sub-type (disabling relay function—two relays disabled & fault)	
35	Fault from 10–49% on L1 with R1 and R2 disabled & fault
36	Fault from 50–90% on L1 with R1 and R2 disabled & fault
37	Fault from 10–49% on L1 with R3 and R4 disabled & fault
38	Fault from 50–90% on L1 with R3 and R4 disabled & fault
Attack sub-type (disabling relay function—two relay disabled & line maintenance)	
39	L1 maintenance with R1 and R2 disabled
40	L1 maintenance with R1 and R2 disabled

4.2 Confusion Matrix

Table 7 shows the confusion matrix for the proposed LWCSO-PKM approach. A confusion matrix contains the information about actual and predicted classifications done by a classification approach. Performance of the classification approach is commonly evaluated using the data in the confusion matrix. True Positive (TP) is the number of correct classification of the suspicious node as an attack node, False Positive (FP) is the number of incorrect classification of the normal node as an attack node, False Negative (FN) is the number of incorrect classification of the attack node as a normal node, and True Negative (TN) is the number of correct classification of the normal nodes. The classification rate is high, if the TP value is high and FP is low. Our proposed LWCSO-PKM approach achieves high TP value and low FP

Table 7 Confusion matrix

Confusion matrix	Predicted	
	Normal	Attack
Actual		
Normal	227	16
Attack	21	736

Table 8 Comparative analysis of various performance metrics for the proposed LWCSO-PKM approach and SVM classifier

Parameters	LWCSO-PKM	SVM classifier
TP	240	179
TN	749	450
FP	3	254
FN	8	117
Sensitivity (%)	96.77	60.47
Specificity (%)	99.6	63.92
Precision (%)	98.7	56.32
Recall (%)	98.9	60.49
Jaccard coefficient (%)	98.3	62.16
Dice coefficient (%)	99.12	76.67
Kappa coefficient (%)	98.65	53.76
Accuracy (%)	98.9	62.16

value. Therefore, the proposed approach achieves efficient classification of the normal and attack nodes.

Table 8 shows the comparative analysis of various performance metrics for the proposed LWCSO-PKM approach and SVM classifier. The TP and TN values of the proposed approach are greater than the SVM classifier. The FP and FN values of the proposed approach are lower than the SVM classifier. The proposed approach achieves high sensitivity, specificity, precision, recall, Jaccard coefficient, Dice coefficient, Kappa coefficient and accuracy than the SVM classifier.

4.3 False Rejection Rate (FRR)

The FRR is defined as the ratio of the number of false rejections to the number of the classified nodes.

$$\text{FRR} = \frac{\text{Number of false rejections}}{\text{Total Number of classified nodes}} \tag{17}$$

The FRR reduces with the increase in the number of classes. Reduction in the FRR of the incorrectly predicted nodes indicates the effectiveness of the classification approach.

Table 9 FRR, FAR and GAR analysis of the proposed LWCSO-PKM approach

Class Label	FRR (%)	FAR (%)	GAR (%)
1	0	3.7838	100
2	3.7838	0	96.2162

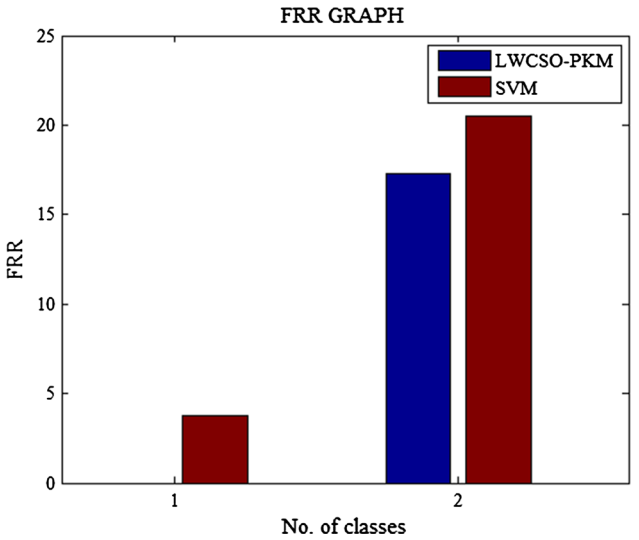

Fig. 2 FRR analysis of the proposed LWCSO-PKM approach

4.4 False Acceptance Rate (FAR)

FAR typically is defined as the ratio of the number of false acceptances to the number of classified nodes.

$$\text{FAR} = \frac{\text{Number of false acceptances}}{\text{Total Number of classified nodes}} \tag{18}$$

The FAR seems to increase with the increase in the number of classes. Hence, the incorrect classification of the nodes is prevented.

4.5 Genuine Acceptance Rate (GAR)

The GAR is the fraction of the genuine scores exceeding the threshold value. Higher the GAR value, higher is the classification efficiency.

$$\text{GAR} = 1 - \text{FRR} \tag{19}$$

Table 9 illustrates the FRR, FAR and GAR analysis of the proposed LWCSO-PKM approach. The proposed approach achieves higher GAR and lower FRR and FAR.

Figure 2 illustrates the FRR analysis of the proposed LWCSO-PKM approach. The FRR of the proposed approach is lower than the SVM classifier.

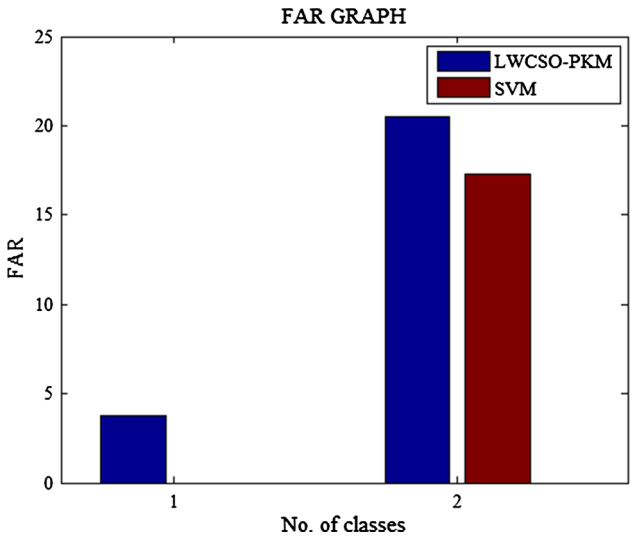

Fig. 3 FAR analysis of the proposed LWCSO-PKM approach

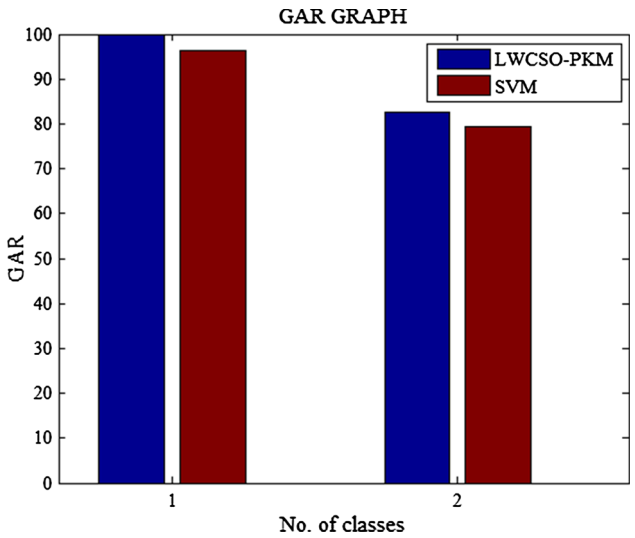

Fig. 4 GAR analysis of the proposed LWCSO-PKM approach

Figure 3 shows the FAR analysis of the proposed LWCSO-PKM approach. The FAR of the proposed approach is higher than the SVM classifier.

Figure 4 shows the GAR analysis of the proposed LWCSO-PKM approach. The GAR of the proposed approach is higher than the SVM classifier. Thus, the classification efficiency of various types of attacks is improved.

4.6 ROC Plot for Classification

The ROC curve is a graphical plot that illustrates the classification performance of the PKM approach. The ROC curve is generated by plotting the comparison of the true positive rate versus the false positive rate at various threshold settings. Each point on the ROC plot represents a pair of the sensitivity/specificity values corresponding to the decision threshold value. The accuracy of the classification approach is high, if the ROC plot is located closer to the upper left corner. Figure 5 shows the ROC curve for classification for the proposed LWCSO-PKM and existing SVM classifier. From the figure, it is clearly evident that the proposed approach achieves higher classification efficiency than the SVM classifier.

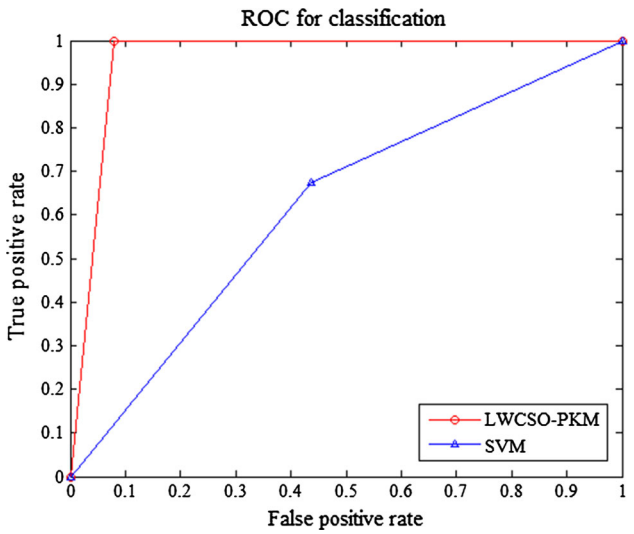

Fig. 5 ROC for classification plot

Table 10 Features selected with their ranges

Feature Name	Ranges
Line current magnitude	(High, Warning, Normal, Zero)
snort_log	(True, False)
relay_log	(True, False)
control_log	(True, False)
Comm_read_function	(3 or 1)
Resp_read_function	(3 or 1)
Control mode	(0 or 1 or 2)
Measurement	6 to 11 or 1 to 100

Table 10 presents the features selected and their ranges during the implementation of proposed work.

Table 11 presents the classification result analysis for negative alarm rate, HH alarm, above H set point, above L set point, LL alarm. The Mississippi State University (MSU) SCADA security laboratory is used for the analysis of the classification result [33] for the water tank control system.

In the water tank control system, the Remote Terminal Unit (RTU) ladder logic program maintains the water level in the tank between the Low (L) and High (H) set points using a controller technique. When it is detected that the water level has reached the L level, the RTU ladder logic program turns

Table 11 Classification result analysis for negative alarm rate, HH alarm, above H set point, above L set point, LL alarm

Parameters	Neural network classifier	Proposed classifier
Negative alarm rate		
False positive rate (%)	0	0
False negative rate (%)	0	0
Accuracy (%)	100	100
HH alarm		
False positive rate (%)	4.5	0.9
False negative rate (%)	0	1.7
Accuracy (%)	95.5	97.4
Above H set point		
False positive rate (%)	2.3	1.2
False negative rate (%)	3	1.7
Accuracy (%)	94.7	97.1
Above L set point		
False positive rate (%)	2.4	1.2
False negative rate (%)	3	1.5
Accuracy (%)	94.6	97.3
LL alarm		
False positive rate (%)	3.2	0.9
False negative rate (%)	0	1.2
Accuracy (%)	96.8	97.9

ON the water pump. If the water level reaches the H level, the water pump is turned OFF.

Negative alarm rate Water level is negative. This is an impossible water level value, though it can be injected.
Highest water level (HH) alarm Water level is above the HH alarm set point. This may be a single false measurement or a group of false measurement.
Above H set point Water level is above the H set point though below the HH alarm set point. This may be a single false measurement or a group of false measurement.
Below L set point Water level is below the L set point though above the LL alarm set point. This may be a single false measurement or a group of false measurement.
Lowest water level (LL) alarm Water level is below the LL alarm set point. This may be a single false measurement or a group of false measurement.

4.7 False Positive Rate (FPR)

The FPR is the ratio of negatives cases that are incorrectly classified as positive. The FPR is defined as the proportion of the attacker nodes that are incorrectly classified as normal nodes by the classifier. This is calculated by using the equation

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{20}$$

4.8 False Negative Rate

The FNR is the percentage of obtaining negative classification results out of the total classification results. It is defined as

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}} \tag{21}$$

The proposed LWCSO-PKM approach achieves very low FPR and FNR and 100% accuracy than the existing neural network classifier for the negative alarm rate. The proposed LWCSO-PKM approach yields low FPR and FNR and higher accuracy than the existing neural network classifier for the HH alarm, above 'H' set point, 'L' set point and LL alarm.

4.9 Accuracy

Accuracy is defined as the ratio of number of correctly classified results to the total number of the classified results. The performance of the classification approach is determined based on the number of samples that are correctly and incorrectly predicted by the classification approach. High accuracy shows the efficiency of the classification approach. It is calculated as

$$\text{Accuracy (\%)} = \frac{\text{Number of correctly classified results}}{\text{Total number of classified results}} \times 100 \tag{22}$$

Table 12 Comparison of the classifier results for the Random Forest, JRip, Adaboost+JRip, mining common path algorithm and proposed LWCSO+PKM approach

Parameters	Random Forest	JRip	Adaboost + JRip	Mining common path algorithm	Proposed LWCSO+PKM
Accuracy (%)	80	90	95	93	98.9
Precision	65	89	99	98	98.7
Recall	20	85	93	95	98.9
F-Measure	28	90	95	96	98.79

4.10 Precision

Precision is defined as the ratio of the number of correct classification results to the total number of classification results. It is calculated as

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{23}$$

4.11 Recall and F-measure

Recall is defined as the ratio of number of correct results to the number of returned results. F-measure is taken as a weighted average of the precision and recall values. Higher values of precision, recall and F-measure indicate the classification efficiency.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{24}$$

$$\text{F1Score} = \frac{2*(\text{Recall} * \text{Precision})}{\text{Recall} + \text{Precision}} \tag{25}$$

4.12 Simulation Scenario

Power System Faults The symmetrical and unsymmetrical faults in a power system are considered as the disturbances. A power system fault is an abnormal condition of the system voltage, current and frequency. The single-line-to-ground (1LG) faults, double LG (2LG) faults, 3LG faults and line-to-line (LL) faults represent more than 95% faults in the power system. The phase-a-to-ground fault for 1LG faults, phase-a-b-to-ground faults for 2LG faults, phase a-b-c-tor-ground fault for 3LG faults and phase-a-to-b LL faults are simulated.
Trip Command Injection Attack These attacks create emergencies through the remote transmission of unexpected relay trip commands from the attacker's computer terminals to the relays located at the ends of a power transmission line.
Aurora Attack It refers to the potential harm caused to a generator due to the intentional rapid opening and closing of a breaker near the generator.
1LG Fault Replay Attack This attack tries to emulate a valid fault by changing the measurements to imitate a 1LG fault and sending an illegal trip command from the attacked computer to the relay. This leads to confusion and causes the control operator to take invalid control actions.

Totally, 1023 1LG fault instances, 274 2LG fault instances, 584 3LG fault instances, 272 LL fault instances, 274 command injection attack instances, 225 aurora attack instances and 703 1LG fault replay attack instances are used as a test data for simulation. The accuracy rate is defined as the percentage of correct classification of the negative instances. The misclassification rate is the percentage of negative instances that are misclassified as positive instances [34]. Table 12 illustrates the comparison of the classifier results for the Random Forest, JRip, Adaboost + JRip [35], mining common path algorithm [34] and proposed LWCSO+PKM approach. Figure 6 depicts the comparative classification results for the Random Forest, JRip, Adaboost + JRip, mining common path algorithm and proposed LWCSO + PKM approach. The Adaboost+JRip achieve highest precision value than the Random Forest, JRip, mining common path algorithm and proposed LWCSO + PKM. The proposed LWCSO + PKM approach yields highest accuracy, recall and F-measure.

4.13 Simulation Scenario

The Australian Defense Force Academy-Linux Dataset (ADFA-LD) [30] is used for performance evaluation. This dataset uses a fully patched Ubuntu Linux 11.04 [32] as the host Operating System (OS). This dataset provides a better indication of performance against the attacks. Table 13 shows the comparison between contemporary IDS algorithms including Data mining of audit files [36], Multivariate statistical analysis of audit data [37], Hidden Markov Model (HMM) and entropy analysis of system calls [38], System call *n*-gram sliding window (assorted decision engines) [39], RBF Artificial Neural Network (ANN) analyzing system calls [40], Multilayer Perceptron (MLP) ANN on subset of KDD98 [41], SVM on subset of KDD98 [41], *k*-Nearest Neighbor (kNN) with Smooth Binary Weighted RBF [42], Rough set clustering [43] and Extreme Learning Machine (ELM) using original semantic feature [30]. The proposed LWCSO-PKM approach yields high detection rate and false alarm rate than the IDS algorithms. From the classification result, if the attack type is already known, then appropriate action is taken to the predicted node. But, if the attack type is

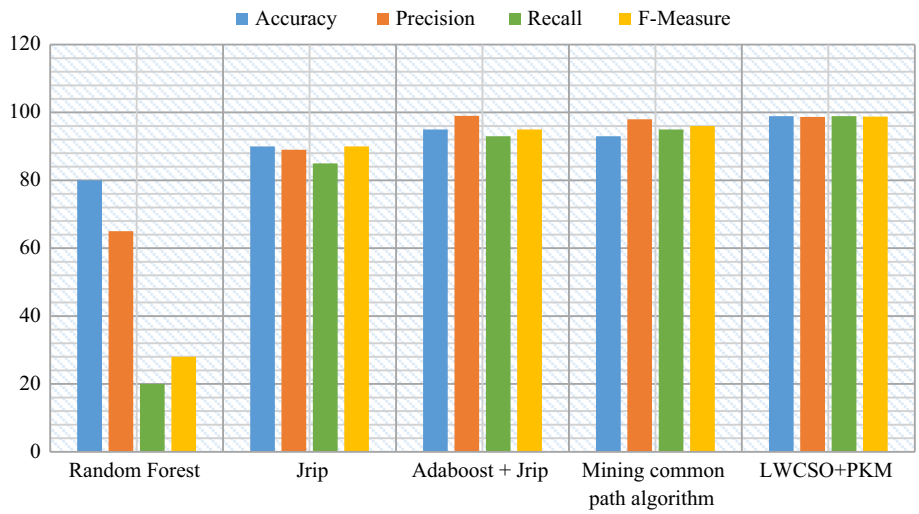

Fig. 6 Comparative classification results for the Random Forest, JRip, Adaboost + JRip, mining common path algorithm and proposed LWCSO + PKM approach

Table 13 Comparison between contemporary IDS algorithms

Algorithm	Detection rate (%)	False alarm rate (%)
Data mining of audit files [36]	80.2	–
Multivariate statistical analysis of audit data [37]	90	40
HMM and entropy analysis of system calls [38]	91.7	10
System call *n*-gram sliding window (assorted decision engines) [39]	96.5	6
RBF ANN analyzing system calls [40]	96 mean	5.4 mean
MLP ANN on subset of KDD98 [41]	99.2	4.94
SVM on subset of KDD98 [41]	99.6	4.17
kNN with Smooth Binary Weighted RBF [42]	96.3	6.2
Rough Set Clustering [43]	95.9	7.2
ELM using original semantic feature [30]	100	0.6
Proposed LWCSO-PKM	100	0.4

unknown, its type is updated to the database and a new label is assigned to it.

4.14 Detection Rate (DR)

The DR is defined as the ratio of the number of detected attacks to the total number of attacks present in the network. It is defined as

$$DR = \frac{\text{Number of detected attacks}}{\text{Total number of attacks}} \times 100 \tag{26}$$

4.15 False Alarm Rate

The false alarm rate (F) is the ratio of the number of false alerts to the number of traces in the validation data. It is calculated by using the following equation

$$F = \frac{\text{Number of false alerts}}{\text{Number of traces invalidation data}} \times 100 \tag{27}$$

Figure 7 shows the ROC for classification analysis plot for the proposed LWCSO-PKM, and existing techniques such as SVM, HMM with syntactic feature, ELM with syntactic feature, single state SVM with syntactic feature and syntactic feature using STIDE length 5. The proposed LWCSO-PKM approach achieves better ROC for classification than existing techniques.

5 Conclusion and Future Work

Our research work proposes a novel model for feature optimization and classification of the attack types in the SCADA network. The memory usage is determined based on the usage of total number of attributes and number of observations.

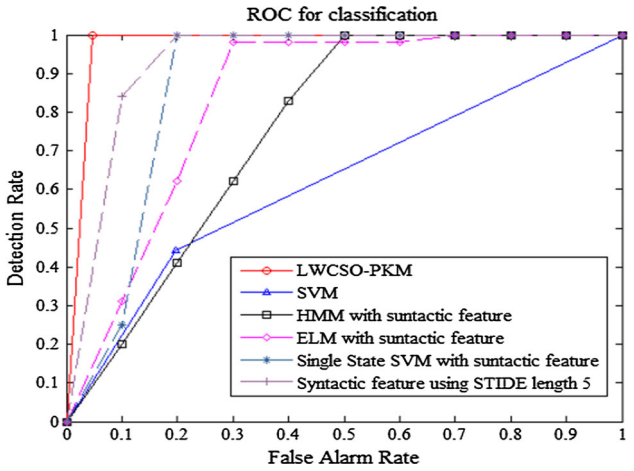

Fig. 7 ROC for classification for the proposed LWCSO-PKM, SVM, HMM, ELM, Single State SVM with suntactic feature and Syntactic feature using STIDE length 5

Due to the architectural constraints and the vulnerability of attacks, high-security based SCADA development is the major requirement. The observations from the sensor play the major role in the classification of normal and abnormal patterns. The high-dimensionality of the features minimized the classification performance. This paper proposed a novel approach for feature optimization and classification of the attack types in the SCADA network with better performance than the existing algorithms. The LWCSO algorithm in proposed work selected the best features from the overall features. A PKM updated the weight function of each node to form the clusters representing the optimal features. The label is applied to each cluster based on the difference between the set of labeled training features with the testing feature set. Based on this label, the features are applied to detect the anomaly node in the network area. From the classification result, if the attack type is already known, then appropriate action is taken immediately. If the attack type is unknown, its type is added to the database. The periodical discovery of the type of attack and the database update with the unknown attacks increased the detection ability effectively. From the performance analysis, it is observed that the proposed LWCSO-PKM approach achieved better performance than the existing classification techniques and IDS algorithms.

References

1. Eesa, A.S.; Orman, Z.; Brifcani, A.M.A.: A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. Expert Syst. Appl. **42**, 2670–2679 (2015)
2. Elhag, S.; Fernández, A.; Bawakid, A.; Alshomrani, S.; Herrera, F.: On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems. Expert Syst. Appl. **42**, 193–202 (2015)
3. Wang, G.; Hao, J.; Ma, J.; Huang, L.: A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering. Expert Syst. Appl. **37**, 6225–6232 (2010)
4. Yang, Y.; McLaughlin, K.; Sezer, S.; Littler, T.; Im, E.G.; Pranggono, B.; et al.: Multiattribute SCADA-specific intrusion detection system for power networks. IEEE Trans. Power Deliv. **29**, 1092–1102 (2014)
5. Zhu, B.; Sastry, S.: SCADA-specific intrusion detection/prevention systems: a survey and taxonomy. In: Proceedings of the 1st Workshop on Secure Control Systems (SCS) (2010)
6. Almalawi, A.; Yu, X.; Tari, Z.; Fahad, A.; Khalil, I.: An unsupervised anomaly-based detection approach for integrity attacks on SCADA systems. Comput. Secur. **46**, 94–110 (2014)
7. Snort_inline. http://snort-inline.sourceforge.net/oldhome.html
8. Yang, Y.; McLaughlin, K.; Littler, T.; Sezer, S.; Pranggono, B.; Wang, H.: Intrusion detection system for IEC 60870-5-104 based SCADA networks. In: IEEE Power and Energy Society General Meeting (PES), vol. 2013, pp. 1–5 (2013)
9. Maglaras, L.A.; Jiang, J.: Ocsvm model combined with k-means recursive clustering for intrusion detection in scada systems. In: 10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness (QShine), pp. 133–134 (2014)
10. Fahad, A.; Tari, Z.; Khalil, I.; Habib, I.; Alnuweiri, H.: Toward an efficient and scalable feature selection approach for internet traffic classification. Comput. Netw. **57**, 2040–2057 (2013)
11. Gong, Y.; Fang, Y.; Liu, L.; Li, J.: Multi-agent intrusion detection system using feature selection approach. In: Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), pp. 528-531 (2014)
12. Nader, P.; Honeine, P.; Beauseroy, P.: I_p-norms in one-class classification for intrusion detection in SCADA systems. IEEE Trans. Ind. Inform. **10**, 2308–2317 (2014)
13. Erez, N.; Wool, A.: Control variable classification, modeling and anomaly detection in Modbus/TCP SCADA systems. Int. J. Crit. Infrastruct. Prot. **10**, 59–70 (2015)
14. Moon, D.; Im, H.; Kim, I.; Park, J.H.: DTB-IDS: an intrusion detection system based on decision tree using behavior analysis for preventing APT attacks. J. Supercomput. (2015). doi:10.1007/s11227-015-1604-8
15. Lin, H.; Slagell, A.; Kalbarczyk, Z.; Sauer, P.W.; Iyer, R.K.: Semantic security analysis of SCADA networks to detect malicious control commands in power grids. In: Proceedings of the First ACM Workshop on Smart Energy Grid Security, pp. 29–34 (2013)
16. Gao, W.; Morris, T.H.: On cyber attacks and signature based intrusion detection for MODBUS based industrial control systems. J. Digit. Forensics Secur. Law JDFSL **9**, 37 (2014)
17. Wang, Y.; Xu, Z.; Zhang, J.; Xu, L.; Wang, H.; Gu, G.: SRID: state relation based intrusion detection for false data injection attacks in SCADA. In: Computer Security-ESORICS 2014. Springer, pp. 401–418 (2014)
18. Rusu, D.A.; Genge, B.; Siaterlis, C.: SPEAR: a systematic approach for connection pattern-based anomaly detection in SCADA systems. Procedia Technol. **12**, 168–173 (2014)
19. Jin, S.; Dan, T.; Zhang, L.; Liu, L.: A fuzzy Bayesian approach to enhance SCADA network security. In: Proceedings of International Conference on Computer Science and Information Technology, pp. 115–122 (2014)
20. Nasr, P.M.; Varjani, A.Y.: Alarm based anomaly detection of insider attacks in SCADA system. In: Smart Grid Conference (SGC), Tehran, Iran, vol. 2014, pp. 1–6 (2014)
21. McLaughlin, K.; Sezer, S.; Smith, P.; Ma, Z.; Skopik, F.: PRECYSE: cyber-attack detection and response for industrial control systems. In: Proceedings of the 2nd International Symposium on ICS and SCADA Cyber Security Research 2014, pp. 67–71 (2014)

22. Jiang, J.; Yasakethu, L.: Anomaly detection via one class svm for protection of scada systems. In: International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Beijing, pp. 82–88 (2013)
23. Do, V.L.; Fillatre, L.; Nikiforov, I.: A statistical method for detecting cyber/physical attacks on SCADA systems. In: IEEE Conference on Control Applications (CCA), pp. 364–369 (2014)
24. Hug, G.; Giampapa, J.A.: Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks. IEEE Trans. Smart Grid **3**, 1362–1370 (2012)
25. Yang, Y.; McLaughlin, K.; Littler, T.; Sezer, S.; Im, E.G.; Yao, Z., et al.: Man-in-the-middle attack test-bed investigating cyber-security vulnerabilities in Smart Grid SCADA systems. In: International Conference on Sustainable Power Generation and Supply (SUPERGEN 2012), pp. 1–8 (2012)
26. Yang, X.-S.; Deb, S.: Engineering optimisation by cuckoo search. Int. J. Math. Model. Numer. Optim. **1**, 330–343 (2010)
27. Yang, X.-S.; Deb, S.: Cuckoo search via Lévy flights. In: World Congress on Nature and Biologically Inspired Computing, 2009: NaBIC 2009, pp. 210–214 (2009)
28. Power System Attack Datasets—Mississippi State University and Oak Ridge National Laboratory—4/15/2014 (2014). http://www.ece.uah.edu/~thm0009/icsdatasets/PowerSystem_Dataset_README.pdf
29. Hsu, J.; Mudd, D.; Thornton, Z.: Mississippi State University Project Report-SCADA Anomaly Detection. http://www.ece.uah.edu/~thm0009/icsdatasets/MSU_SCADA_Final_Report.pdf (2014)
30. Creech, G.; Hu, J.: A semantic approach to host-based intrusion detection systems using contiguousand discontiguous system call patterns. IEEE Trans. Comput. **63**, 807–819 (2014)
31. Creech, G.: The ADFA Intrusion Detection Datasets. https://www.unsw.adfa.edu.au/australian-centre-for-cyber-security/cybersecurity/ADFA-IDS-Datasets/ (2013)
32. Ubuntu. Ubuntu Linux. http://www.ubuntu.com
33. Morris, T.; Srivastava, A.; Reaves, B.; Gao, W.; Pavurapu, K.; Reddi, R.: A control system testbed to validate critical infrastructure protection concepts. Int. J. Crit. Infrastruct. Prot. **4**, 88–103 (2011)
34. Pan, S.; Morris, T.; Adhikari, U.: Classification of disturbances and cyber-attacks in power systems using heterogeneous time-synchronized data. IEEE Trans. Ind. Inform. **11**, 650–662 (2015)
35. Borges Hink, R.C.; Beaver, J.M.; Buckner, M.A.; Morris, T.; Adhikari, U.; Pan, S.: Machine learning for power system disturbance and cyber-attack discrimination. In: 7th International Symposium on Resilient Control Systems (ISRCS), 2014, pp. 1–8 (2014)
36. Lee, W.; Stolfo, S.J.; Mok, K.W.: A data mining framework for building intrusion detection models. In: Proceedings of the 1999 IEEE Symposium on Security and Privacy, pp. 120–132 (1999)
37. Ye, N.; Emran, S.M.; Chen, Q.; Vilbert, S.: Multivariate statistical analysis of audit trails for host-based intrusion detection. IEEE Trans. Comput. **51**, 810–820 (2002)
38. Yeung, D.-Y.; Ding, Y.: Host-based intrusion detection using dynamic and static behavioral models. Pattern Recognit. **36**, 229–243 (2003)
39. Warrender, C.; Forrest, S.; Pearlmutter, B.: Detecting intrusions using system calls: alternative data models. In: Proceedings of the 1999 IEEE Symposium on Security and Privacy, 1999, pp. 133–145 (1999)
40. Ahmed, U.; Masood, A.: Host based intrusion detection using RBF neural networks. In: International Conference on Emerging Technologies, 2009. ICET 2009, pp. 48–51 (2009)
41. Chen, W.-H.; Hsu, S.-H.; Shen, H.-P.: Application of SVM and ANN for intrusion detection. Comput. Oper. Res. **32**, 2617–2634 (2005)
42. Sharma, A.; Pujari, A.K.; Paliwal, K.K.: Intrusion detection using text processing techniques with a kernel based similarity measure. Comput. Secur. **26**, 488–495 (2007)
43. Rawat, S.; Gulati, V.P.; Pujari, A.K.: A fast host-based intrusion detection system using rough set theory. In: Transactions on Rough Sets IV, pp. 144–161. Springer (2005)