I. J. Computer Network and Information Security, 2018, 12, 11-20
Published Online December 2018 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2018.12.02

An Integrated Perceptron Kernel Classifier for Intrusion Detection System

Ruby Sharma [1], Sandeep Chaurasia [2]
[1] Research Scholar, Department of Computer Science Engineering, Manipal University Jaipur
[2] Sandeep Chaurasia, Department of Computer Science Engineering, Manipal University Jaipur
E-mail: study.ruby@gmail.com, sandeep.chaurasia@jaipur.manipal.edu

Received: 08 September 2018; Accepted: 18 October 2018; Published: 08 December 2018

***Abstract*—**Because of the tremendous growth in the network based services as well as the sharing of sensitive data, the network security becomes a challenging task. The major risk in the network is the intrusion. Among various hardening system, intrusion detection system (IDS) plays a significant role in providing network security. Several traditional techniques are utilized for network security but still they lack in providing security. The major drawbacks of these network security algorithms are inaccurate classification results, increased false alarm rate, etc. to avoid these issues, an Integrated Perceptron Kernel Classifier is proposed in this work. The input raw data are preprocessed initially for the purpose of removing the noisy data as well as irrelevant data. Then the features form the preprocessed data are extracted by clustering it depending up on the Fuzzy C-Mean Clustering. Then the clustered features are extracted by employing the Density based Distance Maximization approach. After this the best features are selected using Modified Ant Colony Optimization by improving the convergence time. Finally the extracted best features are classified for identifying the network traffic as normal and abnormal by introducing an Integrated Perceptron Kernel Classifier. The performance of this framework is evaluated and compared with the existing classifiers such as SVM and PNN. The results prove the superiority of this framework with better classification accuracy.

***Index Terms*—**Network Security, Intrusion Detection System, Fuzzy C-Means Clustering, Density based Distance Maximization approach, Ant Colony Optimization, Ensemble Classifier.

I. INTRODUCTION

With the arrival of the internet and innovative technology in networking, the world has become more interlinked. This networking infrastructure has large amount of information that are related to the commercial, government, military and personal [1]. Due to this widespread utility of internet and networking applications, there arise some security issues in which the attackers tried to hack the information in the network. In order to protect the information, it is essential to secure the network. Network security plays a significant role due to its intellect which can be acquired easily through the internet. It is the process of protecting the system from an unauthorized access. The network security comprises the issues such as maintenance of integrity in the network, prevention of authorized users and the protection of network. When the size of the network becomes increased, then these issues are magnified to a great extent [2]. In current situation, it is a crucial task for implementing the security methods. Generally there are several attacks which destroy or destructing the connections in the network and also the communication between them. These attacks may cause some limitations such as slow network performance, uncontrolled traffic, etc.

Most commonly the attacks can be classified into two broad groups such as active attack and passive attack. In which an attacker may initiate the commands for disrupting the normal operation of the network referred to as active attack. This type of active attacks has some examples such as wormhole, spoofing, sinkhole, denial of service, Sybil attack and modification. Whereas, a network attacker may intercept the transmission of data through the network is referred to as passive attack. Some examples of this passive attack are represented as eavesdropping, traffic analysis, monitoring, etc. Also the network system has some advanced attacks such as black hole attack, byzantine attack, replay attack, location disclosure attack and rushing attack [3]. Several traditional methods are utilized for securing the network from the attackers. Some of them are intrusion detection systems, firewalls and authentication devices. From these various security methods, intrusion detection system plays a significant part for protecting the networks. In earlier days, the issues in the intrusion detection system received greater attention. Various traditional works such as data mining, statistical analysis, machine learning techniques and immunological inspired methods.

Generally IDS are categorized as three types such as Signature based, Anomaly based and Hybrid IDS [4]. In this signature based IDS system, the behavior of the users is compared with the known activities of the attackers for the purpose of penetrating a network or a system. In this system, the collected data is analyzed and compared with the large database for attack signatures. It helps to detect the known attacks in an effective manner with less false alarms. But it is hard to detect the zero day attacks in this

approach. In anomaly based IDS system, the behaviors which are varied from the users' pattern that are established can be detected. For the normal behaviors of the users, a profile is established. This profile can be compared with the data of the actual users when the detection process is performed. When the threshold value is greater than the offset value, the behavior of the user can be deliberated as normal. Whereas, the threshold value is lesser than the offset values, then the behavior of the user can be deliberated as abnormal or malicious. They have the capability to detect the zero day attacks and thus it is interesting. It is hard to know the activities which can be carried out as undetected since the profiles of the normal behaviors are personalized for each system or network. Moreover it has the major drawback of high false alarms due to the categorization of the behavior of earlier undetected system as anomalies. Another type of IDS technique which is referred to as Hybrid IDS technique integrates both the signature and anomaly detection. This hybrid IDS technique helps to increase the detection rate of the known attacks and also reduces the false alarm of the unknown attacks.

Furthermore, this IDS technique can be partitioned into two major groups depending up on the intrusive behavior such as network based IDS technique and host based IDS technique. The network based IDS technique detects the intrusions by monitoring the traffic through the network devices. Whereas, the host based IDS technique detect the activities of the file or process which are related to the software environment that are linked with a particular host. Several existing techniques have been utilized for detecting the anomalies and intrusions in the network. Although they have several major benefits such as better detection rates, less false alarm, etc. but still it lacks in the issues like inaccurate detection and classification results. Also there are many attacks in the network due to the issues of respective layers. Thus the optimization techniques are utilized for obtaining better results in IDS. Hence this work is aimed to determine the performance of a novel intrusion detection in both with and without optimization techniques. Thus a novel detection framework is proposed for intrusion detection system to determine the better performances.

The main objectives of this work are listed as follows:

- To create an effective IDS system using an Integrated Perceptron Kernel Classifier based Intrusion Detection Framework.
- To extract the features from the preprocessed data using Fuzzy C-Means clustering.
- To obtain the best features by utilizing Modified Ant Colony Optimization technique.
- To classify the network traffic using an Improved Ensemble classifier model.

The remaining sections of this paper are systemized as follows: Section II reviews various existing techniques that are related to the detection and classification of IDS along with its advantages and disadvantages. Section III gives the detailed descriptions about the flow of proposed methodology. Section IV evaluates the performance of the projected approach. Finally the work is concluded in section V.

II. Related Works

This section reviews the various techniques that are utilized for intrusion detection systems. Also it discusses the advantages and disadvantages of the traditional techniques that help to detect the intrusions in the systems. [5] considered the utilization of genetic fuzzy systems in a framework which is based on the pairwise learning for the progress of those system. This approach had two benefits: one was the utilization of fuzzy sets and particularly the linguistic labels enabled a smoother borderline amongst the perceptions and permitted an increased interoperability of the rule set. Another was the learning schemed named divide and conquer wherein all the probable pair of classes were contrasted with the objectives enhanced the precision of the events of rare attacks since it obtained best separability amongst a normal actions and various types of attacks. Initially this approach helped to manage the numerical variables that were connected with the intrusion detection in a better way. Also the strategy of divide and conquer enhanced the classification accuracy. But still it had lacked in some evaluation metrics and it did not have the ability to select the best solutions. [6] proposed a novel fuzziness based semi supervised learning approach by the utilization of unlabeled samples which were assisted with the supervised learning algorithm. The main intention of this approach was to enhance the performance of the classifier for the intrusion detection systems. A fuzzy membership vector could be obtained as an output by training a single hidden layer feed forward neural network and the categorization of the samples could be performed on unlabeled samples by utilizing the fuzzy quantity. After integrating each category into original training set, the classifier was retrained in a separate manner. From the results it was observed that the samples which were belonged to high and low fuzziness played a significant part to enhance the performance of the classifier. But the disadvantage was that this strategy was nor implemented in the detection of multiple attacks.

[7] introduced a noel ensemble construction technique which utilized the weights that were generated by PSO for creating the ensemble of classifiers with increased intrusion detection accuracy. For PSO, the best behavioral parameters were found by utilizing Local Unimodal Sampling (LUS) as a Meta optimizer. The better accuracy could be gained by the LUS but it had increased running time. Hence it was essential to implement the technique with different optimization approaches to generate the weights. [8] presented the development and performance evaluation of a host based misuse detection system. Here an ensemble design was employed for the classification and the raw call trace data of the ADFA-LD system was preprocessed by N gram feature extraction method. In this method, the dataset had the modelling of six different attacks. The dimensionality of the input patterns was reduced and the unique signature of each class was

captured by two filters. Also the unique features from each class which were most frequently in the form of N grams were extracted. The design of classifier was depending up on the majority voting ensemble. This evaluation system resulted an increased performance for the detection of the attacks in the issues of binary class. But still it was lacked in the detection of intrusion in multi class systems.

[9] suggested a novel feature representation method which was referred to as cluster center and nearest neighbor method (CANN). Here, there were two distances in which, one was the distance amongst the sample of each data and the cluster center of those sample and the other was the distance amongst the data and the nearest neighbor that data in the same cluster. These two distances were measured and summed which was resulted as a one dimensional distance. This one dimensional distance based feature was utilized for representing each data sample by utilizing a K nearest neighbor classifier for the purpose of detection the intrusions. The results showed that the suggested CANN classifier performed better than other classifiers and also it offered an increased efficiency of the computing the detection. But this CANN had the drawback of ineffective detection of some attacks such as U2L and R2L. [10] determined a new combination approach depending up on the ID3 algorithm and the bees algorithm for selecting the best subset of features for IDS. The subset of features was generated by utilizing the bees algorithm and the ID3 algorithm was utilized as a classifier. The experimental results revealed that the feature subset that was generated by the determined ID3-BA became superior in the metrics such as detection rate and accuracy rate. Also it had less false acceptance rate. But it was ineffective for the neighborhood search since it had uncontrolled number of bees.

[11] aimed to detect the significant features for the intrusion detection system in an effective and computationally efficient manner. An intrusion detection system was proposed in which the optimal features were selected by utilizing ant colony optimization. Because of the utilization of a simplified feature set for the classification purpose, the proposed method was implemented in an easy manner and also it had less computational complexity. This method reduced the memory size and the utilization of number of features for detecting the intrusions. This in turn reduced the CPU time which was required for intrusion detection. But it was not suited for the intrusion detection based on the packet payload. And further it had poor detection rate. [12] introduced an anomaly detection system at the hypervisor layer which was referred to as Hypervisor detector. It utilized a hybrid algorithm which was the combination of fuzzy c means clustering and artificial neural network for enhancing the detection accuracy of the system. The performance of this introduced technique proved its superiority in classification when compared to the existing classifiers. It also offered better classification results even for the low frequent attacks. This hypervisor detector had increased detection rate and less false alarm rate for the detection of several attacks.

[13] presented a new feature selection approach for the intrusion detection systems. This approach was mainly depending up on the cuttlefish algorithm. The main purpose of utilizing a new feature selection approach was because of the large number of features were utilized in IDS. Also the main intention of this approach was to determine the features with best quality. The cuttlefish algorithm was utilized as a searching strategy for ascertaining the best subset of features and the decision of the selected features were obtained by utilizing the decision tree classifier. The evaluation of the presented feature selection approach showed that the feature subset which was obtained by cuttlefish algorithm offered an increased detection rate as well as accuracy rate with reduced false alarm rate. [14] introduced a novel support vector machine (SVM) model combined with the kernel principal component analysis (KPCA) with genetic algorithm (GA) for detecting intrusions in the system. In order to estimate the attacks in an activity, the multi-layer SVM classifier was adopted and the dimensions of the feature vectors could be reduced as well as training time would be shortened by utilizing KPCA as a preprocessor of SVM. The noise which was caused by the differences in the features was reduced and enhance the performance of SVM using an improved kernel function. Here the mean value and the difference value of the mean square of the feature attributes in the RBF kernel function was embedded. The kernel parameters, punishment factor and tube size of SVM were optimized by employing the genetic algorithm. This proposed model gave higher prediction accuracy with high speed of convergence than the other detection techniques.

[15] developed a novel SVM in an integration with the kernel principal component analysis along with an improved chaotic particle swarm optimization for dealing with the detection of intrusion. Here the punishment factors, tube size of SVM and kernel parameters were optimized by utilizing the improved chaotic particle swarm optimization approach in which the chaos optimization and the premature processing approaches were introduced. The main intention of introducing ICPSO was to select appropriate parameters for the SVM classifier. The results showed that the developed KPCA-ICPSO-SVM model performed superior to that of the SVM classifiers and this SVM classifier with the utilization of KPCA attained the performance of better generalization by extracting the features. This was because of the reason that the KPCA explored the information with higher order of the inputs. The performance analysis gave excellent performance in the detection of intrusions and also it save more computational time required for both training and testing. But in case of predicting the attacks in an accurate manner the developed approach was failed particularly in the attacks such as R2L and U2L. [16] focused on the two stage approach for selecting the features depending up on the random forest technique. In the initial stage, the features with higher variable significance score were selected and the initialization of the search process was guided to the second stage in which the final subset of features was obtained as an output for interpretation and classification. From the results it was

noted that the intended approach which was depending up on the random forest technique had the capability to select the most significant and relevant features that were useful for the process of classification. It also reduced the number of input features and computational time. The approach offered better classification results with increased accuracy rate.

[17] anticipated a method for selecting the features in the intrusion detection system depending up on the particle swarm optimization technique. The optimal sunset of the features were selected from the principal component analysis space. The evaluation of the intrusion detection was carried out by testing the performance of the anticipated methodology. Here the PSO based feature subset was validated by utilizing the modular neural network and these features were compared with the feature subsets that were obtained using tradition optimization techniques. The results indicated that the PSO based feature selection approach outperformed compared to the existing techniques. [18] implemented a hybrid method of support vector machine and genetic algorithm and it explained the issues related to the intrusion detection. The main motive of this hybrid algorithm was to reduce the number of features. Here the features were characterized into three priorities by utilizing GA approach. The prioritization could be done as the most important and relevant features were considered as the first priority. The middle important features were considered as the second priority whereas the least significant features were considered as the third priority. The results showed that the hybrid approach had the ability to achieve better true positive and false positive rates.

[19] developed a combined classifier model which was depending up on the tree based algorithm for detecting the network intrusions. The aim of this detection algorithm was to classify the network traffics as normal or abnormal. This was done depending up on the features which described each pattern of the network traffic. While combining the tree algorithms such as random tree and NB tree, the accuracy was increased depending up on the scheme called sum rule and the individual random tree algorithm was outperformed. This resulted that combined classifier approach offered better classification results than the individual classifiers. [20] presented a classification approach which hybridized the statistical approaches and SOM for detecting the anomalies in the network. In this approach, the principal component analysis and the fisher discriminant ratio were deliberated for the purpose of selecting the features and noise removal. Here the feature space was modeled by utilizing the probabilistic self organizing maps and enabled discrimination amongst the normal and anomalous links. The abilities of the detection could be altered by modifying the probability of unit activation without retraining the map. This approach made fast implementation of the intrusion detection systems which further revealed that the computational time required for detecting the intrusions or anomalies in the network became less.

[21] utilized an intelligent system for maximizing the detection rate of the attacks in network. This was carried out by embedding the temporal activities of the attacks into a structure of neural network which was referred to as TDNN. In this system the principal component based neural network was utilized for detecting the attacks in the network and these attacks were classified as port scan and host sweep by utilizing classification module. This utilized system offered better results in detection rate as well as throughput.

III. Proposed Methodology

This section demonstrates the working procedure of the Integrated Perceptron Kernel Classifier approach. Fig. 1 describes the diagrammatic representation of the flow of the proposed technique.

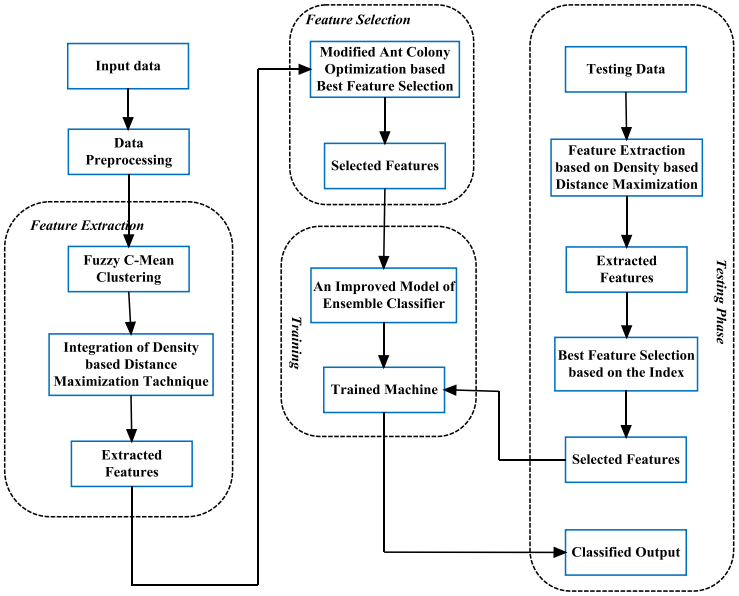


Fig.1. Flow of the Novel Intrusion Detection Framework

Initially the raw dataset is given as the input and pre-processed effectively to remove the noisy data from the network traffic data. Followed by the process of feature extraction is carried out in which the dataset is clustered based on the FCM clustering by the employment of Density based distance maximization Technique. From the clustered data the features can be extracted and the best features are selected based on the Modified Ant colony optimization Technique by improving the convergence time. From the extracted best features, an Improved Model of Ensemble classifier is proposed to classify the network traffic into their corresponding types as Normal and anomaly along with the class types. The symbols used for the algorithms are described in table 1.

Table 1. Definition for Symbols used

Symbols	Description
r_{nd}	Random number
$An_{d(i,j)}$	Ant values
C_{cen}	Center point
Dt	Distance
ρ	Density
Ψ	*cut off distance*
er_c	clustered data
v_f	velocity of the ant
obj_n	*Objective function*
f_{fn}	*Fitness value*
N_c	number of classifier
N_r	number of runs
N_s	number of samples
N_f	number of features
Lb	labels for the respective selected features
D_{pro}	*decision profile*
R_{sort}	*rank level*

A. Feature Extraction using Density based Distance Maximization Technique

A clustering approach can be utilized to reduce the objective function that fits to the clusters objective function algorithms. Here the fuzzy c-mean clustering approach is utilized which is aimed to segment the data into various clusters according to the values of degree membership. Each data may be comprised of data that are similar to each other but different from the data which is from the other cluster.

In this approach, let us consider N_c as the number of clusters and create the default option D_{op}. Then the options for exponential, maximum iteration and minimum improvement are declared. Now let us assume that Exp_c be the exponent for partition m*n files, Min_{imp} be the minimum improvement of the N_c and Max_{ite} be the maximum iteration of N_c. Initially the fuzzy partitions are initialized and the membership functions can be obtained as,

$$mm_{fn} = f_{par}^{Exp_c} \tag{1}$$

After this the center point using this membership function with the exponent for partition files are calculated by the following equation as,

$$C_{cen} = mm_{fn} * Exp_c \Big/ 2\prod\sqrt{mm_{fn} * Exp_c} \tag{2}$$

Then, the distance for each iteration can be calculated using the equation,

$$Dt_{dc} = Exp_c + \left(Max_{ite}, \min_{imp} \Big/ \sqrt{(2\prod((mean(D) * \sqrt{(Deviation(D))^2}}\right) \tag{3}$$

Based on this distance calculation the features are extracted. The algorithm for the feature extraction based on the density based distance maximization technique is given as follows.

Algorithm: I Density based Distance Maximization Technique

Input: *pre-processed m*n files* (Pre_f)
Output: *extracted clusters* Ex_c

Initialize the fuzzy partitions f_{par}
Calculate the membership function using equation (1)
Find the center point by using the equation (2)
Add Max_{ite} *to* f_{par} *// cluster count*
Add Max_{ite} *to* ch_{cen}
For i=1 to Max_{ite}
  For j=1 to J
    $D_{I,J}$ ←*Distance* (Max_{ite}, min_{imp})
Calculate the distance for each iteration using equation (3)
    I←*index* ($min(D_{I,J})$)
  Update mm_{fn} *to* Exp_c
  End
End
Let density ρ_{de}=$\sum_i Dt_{dc}$
Δ_{dt}=*min*(ρ_{de})
Ψ=*cut off distance*
ρ_{deh}=Ψ*Δ_{dt}
ρ_{nor}=*sort*(ρ_{de})
For i=2: ρ_{de}
  $\Delta_{dt}(\rho_{nor})$=*max*($Dt_{dc}(\rho_{nor})$);
For j=1: i-1
  If $Dt_{dc}(\rho_{nor})$< $\Delta_{dt}(\rho_{nor})$
    $\Delta_{dt}(\rho_{nor})$= $Dt_{dc}(\rho_{nor})$
    $\rho_{deh}(\rho_{nor})$= ρ_{nor}
End
End
  End
¥$_{cl}$=*zeros*(1, ρ_{de})

If $C_{cen}(\rho_{nor})$==0
  $Ex_c(\rho_{nor})$=$Ex_c(\rho_{deh}(\rho_{nor}))$
else
  $Ex_c(\rho_{nor})$=$Ex_c((\rho_{nor})$
End

B. Feature Selection using Modified Ant Colony Optimization

The most commonly utilized step in machine learning is the Feature Selection particularly in the high dimensional feature space. The motive of feature selection is to simplify a data by minimizing its dimensionality and recognizing the relevant features. By this, the redundancy in the information is also gets decreased. Due to the issue in the practical applications, the process of feature selection is essential because of the presence of irrelevant features and noisy data. Here the ant colony optimization is utilized which proves its superiority than the traditional optimization techniques. This ACO is inspired the activities of the ant colonies. Generally the ACO can be implemented for the travelling salesman problem. It is specifically attractive for the selecting the features since it leads the search to optimal subset.

In this approach, the modified ant colony optimization is utilized in which the velocity of the ants is determined for updating the ant values. Initially, the clustered data (er_c) are initialized. Let us assume S be the size of the clustered data and An_d be the ant values. Calculate the ant values for each clustered data using the equation,

$$An_{d(i,j)} = \left(\sqrt{er_{ci}} - \sqrt{er_{cj}}\right)^2 - \left(\sqrt{er_{cj}} - \sqrt{er_{ci}}\right)^2 \quad (4)$$

Then the velocity of the ant can be initialized and select the random number r_{nd}. Compute p_s and v_f by using the equation,

$$P_s = (r_{nd(i,k)} + r_{nd(i,k+1)}) - ((r_{nd(i,k+1)} + r_{nd(i,k+1)}) + r_{nd}) \quad (5)$$

$$v_f = \omega(r_{nd} - 1) * v_f(r_{nd}) - An_d \Big/ P_s\left(\sqrt{r_{nd}}\right) \quad (6)$$

Finally update the ant values and determine the fitness value to obtain the optimal features. The algorithm for modified ant colony optimization algorithm is illustrated as follows:

Algorithm II: Modified Ant Colony Optimization

Input: *Extracted Clusters Sequences* Ex_c
Output: *Optimization data* Op_d

S be the Size of the Total sequences to be clustered
Initialize clustered data (er_c)
Extracted clustered from the sequences
Let S= size of (er_c)
Let An_d *be the ant values*
For i=1 to S
  For j=1 to An_d
    Let input data=$er_{c\,i,j}$
  Calculate the ant values for each clustered data using equation (4)
    End
  Let $v_f = x_i - x_j / t$ *//initialize ant velocity*
      Let r_{nd} *= random numbers select*
  Compute p_s*=function(*An_d, r_{nd} *)*
  For i=1 to 255 // random numbers
    For k=1-$r_{nd} - 1$
      Compute p_s v_f *using the equations (5) and (6)*
End
End
Update An_d *values,*
Objective function obj_n
$p_s = obj_n(An_d)$
Fitness value f_{fn}
If $f_{fn} < p_s$
$p_s = f_{fn}$
End

C. Classification using an Integrated Perceptron Kernel Classifier

Generally the classifier is utilized for classifying the system or a network as a normal or abnormal to identify the type of intrusion. Several classifiers are utilized in the conventional techniques. In order to improve the performance of the classification, a novel ensemble classifier is introduced in this work. Here the process of classification is carried out by combining three classification approaches such as Multi-Layer Perceptron (MLP), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). The Multilayer Perceptron classifier is a learning machine with feed forward networks which is comprised of more than two layers. These layers are linked with each other by utilizing the linkage weights and also has a link with the balancing node referred to as bias node. The most commonly as well widely utilized function is the MLP. It has an advantage of easy to maintain the networks. It has the capability to implement on large issues and also it helps to produce non-dependencies.

The K-Nearest Neighbor classifier is a theoretically mature data mining algorithm which has low complexity. The fundamental concept of this KNN is that in a sample space when most of KNN samples are belonging to a group, then that sample belonged to the group. Here the single dimensional or multi-dimensional feature vectors are referred to as the nearest neighbor which can be utilized for describing the sample on the nearest. Also, the nearest criteria is the Euclidean distance of the feature vector. The most familiar machine learning approach is the Support Vector Machine which is employed for solving the issues related to the regression and classification. The fundamental concept to SVM is to determine the feature space of optimal linear hyperplane which maximally splits the two classes that are targeted. The basic SVM utilized a set of input data for the purpose of prediction. For every input that was given which is comprised of two possible classes generates the output or the binary linear classifier. By combining all these algorithms an integrated perceptron kernel classifier is utilized for classifying the intrusions in the system.

Let us consider, N_c be the number of classifier, N_r be the number of runs, N_s and N_f be the number of samples and features respectively. Initially the size of the training

set and the testing set are initialized. Then the features from the sequences are extracted and created a list of feature set. Now assume that, f_{fn} as the extracted feature set, Lb_s as the labels for the respective selected features and N_c as the number of classes to be identified.

The algorithm for an integrated perceptron kernel classifier is discussed as follows:

Algorithm III: Integrated Perceptron Kernel Classifier

Input: *extracted Clusters Sequences* f_{fn}
Output: *Classified output*

Load f_{fn} *//load optimized data*
For i=1 to N_c
  Split F_{fs} *(feature set) into* τ *(feature subset)*
  For j=1 to τ
Switch case 1
Let f_{fn} *is to be feature in* $F_{fs(i,j)}$
    Remove the subset values τ *from* F_{fs}
      Let T_{CLF} *train* (f_{fn}, Lb_s)
        Train out T_{out}=*sim*(T_{CLF}, T_{in})
          D_{pro}=*mapminmax*($T_{out,1}$) //
decision profile
          R_{sort}=*sort*(T_{out}) *//rank level*
  Let compute accuracy
Switch case 2
    Compute Trained T_{CLF}
Let R_{sort}=*sort*(T_{out}) *//rank level*
Switch case 3
  Trained T_{CLF} *estimates* f_{fn}
Let R_{sort}=*sort*(T_{out}) *//rank level*
Let accuracy = mean(T_{CLF} ,1)
  T_{CLF} =$\sum_i T_{out}/ R_{sort}^2$
  $cnt_i = \Sigma(f_{fn})$ *in belonging to samples* N_s
End
Compute Total count as $Cnt_T = \sum_{i=1}^{N} cnt_i$
Compute probabilistic Components for each class as
For i=1 to N_c
  $P_{comp}(i) = cnt_i/Cnt_T$
End

IV. Performance Analysis

This section demonstrates the performance analysis of the proposed framework. It validates the performance compared with the existing techniques. The datasets such as KDD [22] and ADFA [23] are utilized in this work for analyzing the performance of the anticipated framework. This performance analysis is carried out for the suggested methodology by utilizing ACO and compared the results without optimization technique. Here several existing techniques such as SVM and PNN are used for comparing the results.

A. Performance Measures

The performance of the anticipated methodology is evaluated using various performance metrics such as sensitivity, specificity, accuracy, precision, recall, Jaccard, Dice and Kappa coefficients.

Sensitivity

Sensitivity is referred to as the measure of the ratio of True Positive that are recognized accurately. It can be expressed as

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

Specificity

Specificity is described as the ratio of True Negatives that are detected correctly. This can be estimated as

$$Specificity = \frac{TN}{TN + FP} \tag{8}$$

Accuracy

An accuracy is defined as the proportion of correctly classification intrusions to the total number of data. It is explained as,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

Precision

Precision is the ratio of the number of intrusions that are correctly classified to the total number of intrusions in the system. This can be denoted by,

$$\Pr ecision = \frac{TP}{TP + FP} \tag{10}$$

Recall

Recall is the ratio of the number of intrusions that are correctly classified to the number of intrusions that are relevant. This can be represented as,

$$\mathrm{Re}call = \frac{TP}{TP + FN} \tag{11}$$

Jaccard Coefficient

The Jaccard coefficient is characterized as the ratio of intersection of two sets to the union of the two sets. It is calculated as

$$JaccardCoefficient = \frac{A \cup B}{A \cap B} \tag{12}$$

Dice Coefficient

Dice is a mean overlap which determines the

intersection amongst the two sets. It can be calculated as,

$$Dice\ Coefficient = 2\frac{|A \cap B|}{|A| + |B|} \quad (13)$$

Kappa Coefficient

Kappa coefficient is a statistic measure which is utilized for classifying the number of items as categories. It is ranges from -1 to +1. It can be stated as,

$$Kappa\ Coefficient = \frac{R_0 - R_e}{1 - P_e} \quad (14)$$

Where, R_o is the comparative detected agreement amongst the raters, R_e is the theoretical probability of unplanned agreement and P_e is the probability of each observer

B. Performance Analysis

The performance of the suggested framework is evaluated and compared with existing techniques. The comparative results of the anticipated and existing techniques for various parametric measures are shown as follows:

Table 2. Performance analysis of KDD dataset with ACO

Measures	Ensembler	SVM	Na ïve Bayes
Sensitivity	99.8987	74.6891	91.5179
Specificity	99.9857	99.9517	98.9623
Accuracy	99.9714	99.9125	99.9000
Precision	95.9184	72.9602	92.3485
Recall	99.8987	74.6891	91.5179
Dice Coefficient	97.5682	73.7553	91.7358
Jaccard Coefficient	95.8171	72.6530	85.8207
Kappa Coefficient	0.9755	0.7372	0.9118

Table 2 shows the comparative analysis of the ensemble classifier and existing classifiers for the KDD dataset using optimization technique. The results show that the suggested frameworks offers 0.06% increased accuracy results than the traditional SVM approach. Thus the proposed Ensembler classifier detects the intrusion in the network very accurately.

Table 3. Performance analysis of KDD dataset without ACO

Measures	Ensembler	SVM	Na ïve Bayes
Sensitivity	92.4892	79.0476	69.8017
Specificity	99.7349	99.7097	99.7410
Accuracy	99.6222	99.4857	99.5750
Precision	97.9659	83.6793	73.1879
Recall	92.4892	79.0476	69.8017
Dice Coefficient	94.6622	80.7531	71.3584
Jaccard Coefficient	90.6444	77.0126	68.2499
Kappa Coefficient	0.9440	0.8048	0.7112

Table 3 shows the comparative analysis of the ensemble classifier and existing classifiers for the KDD dataset without optimization technique. The results shows that the suggested frameworks offers 0.05% increased accuracy results than the traditional SVM approach. It is observed that the anticipated approach offers better results with optimization than without using optimization.

Table 4. Performance analysis of ADFA dataset with ACO

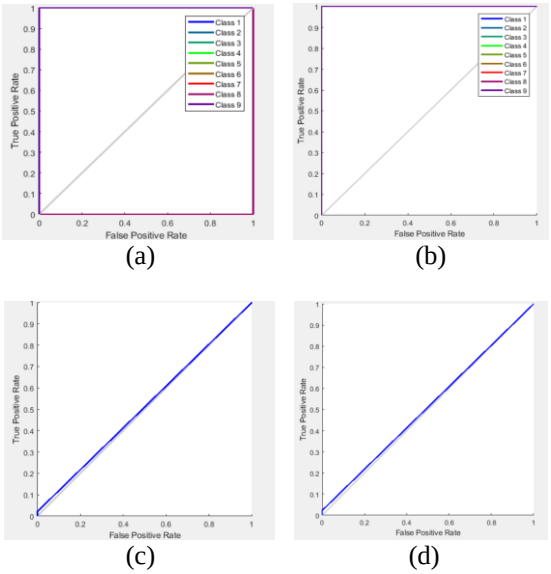
Measures	Ensembler	SVM	Na ïve Bayes
Sensitivity	99.8994	74.8675	68.2398
Specificity	99.9857	99.9667	99.6809
Accuracy	99.9714	99.9375	99.4750
Precision	91.0714	73.3488	67.2893
Recall	99.8998	74.8675	68.2398
Dice Coefficient	94.2352	74.0560	67.5643
Jaccard Coefficient	90.9708	73.2182	63.6209
Kappa Coefficient	0.9422	0.7403	0.6722

Table 4 demonstrates the comparative analysis of the ensemble classifier and existing classifiers for the ADFA dataset using optimization technique. The results shows that the suggested frameworks offers 0.03% increased accuracy results than the traditional SVM approach. Thus the proposed Ensembler classifier detects the intrusion in the network very accurately.

Table 5. Performance analysis of ADFA dataset with ACO

Measures	Ensembler	SVM	Na ïve Bayes
Sensitivity	81.5350	67.2447	59.2882
Specificity	98.6746	98.3120	99.8834
Accuracy	99.8250	99.7250	99.8000
Precision	82.8445	66.0588	58.7665
Recall	81.5350	67.2447	59.2882
Dice Coefficient	73.9331	62.4244	59.0183
Jaccard Coefficient	0.8039	0.6237	58.1036
Kappa Coefficient	0.9422	0.7403	0.5892

Table 5 shows the comparative analysis of the ensemble classifier and existing classifiers for the ADFA dataset without optimization technique. The results shows that the suggested frameworks offers 0.1% increased accuracy results than the traditional SVM approach. It is observed that the anticipated approach offers better results with optimization than without using optimization.



(a) (b) (c) (d)

Fig.2. ROC for two Different Dataset

ROC plot can be measured between true positive rate and false positive rate. The proposed technique is linearly increased from 0 to 1.Fig. 2 (a) denotes the ROC plot for ADFA dataset with ACO, Fig. 2(b) represents the ROC plot for ADFA dataset without ACO, Fig. 2 (c) depicts the ROC plot for KDD dataset with ACO and Fig. 2 (d) shows the ROC plot for KDD dataset without ACO.

V. Conclusion

In recent trends, the services based on the network and sharing the information through network has tremendous growth. Because of this, there may be chance for the existence of intrusions in the network. In order to harden the systems against intrusion, network security is the most important aspect. The traditional techniques are utilized for protecting the information in the network. Among these, intrusion detection system is the most significant methods. Even though, they had several advantages in detecting the intrusions. Still it has some issues such as inaccurate classification results, increased false alarm rate, etc. Thus, an Integrated Perceptron Kernel Classifier is proposed to overcome the above issues in this work. Initially the input raw data are preprocessed by removing the noisy data as well as irrelevant data. Then the features from the preprocessed data are extracted by clustering it depending up on the Fuzzy C-Mean Clustering. Then the clustered features are extracted by employing the Density based Distance Maximization approach. Hereafter, the process of selecting the best features is carried out using Modified Ant Colony Optimization by improving the convergence time. Finally the extracted best features are classified for identifying the network traffic as normal and abnormal by introducing an Integrated Perceptron Kernel Classifier. This integrated classifier combines three classifiers such as MLP, KNN and SVM to obtain the best classification results. The performance of this Integrated Perceptron Kernel Classifier is evaluated and compared with the existing classifiers such as SVM and PNN. From the experimental analysis, it is concluded that the anticipated framework offers superior results with better classification results.

References

[1] B. Daya, "Network security: History, importance, and future," *University of Florida Department of Electrical and Computer Engineering*, vol. 4, 2013.

[2] K. Purohit, "Introduction to Computer Network with Security."

[3] M. V. Pawar and J. Anuradha, "Network security and types of attacks in network," *Procedia Computer Science*, vol. 48, pp. 503-506, 2015.

[4] H. A. M. Uppal, M. Javed, and M. Arshad, "An overview of intrusion detection system (IDS) along with its commonly used techniques and classifications," *International Journal of Computer Science and Telecommunications*, vol. 5, pp. 20-24, 2014.

[5] S. Elhag, A. Fernández, A. Bawakid, S. Alshomrani, and F. Herrera, "On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on intrusion detection systems," *Expert Systems with Applications*, vol. 42, pp. 193-202, 2015.

[6] R. A. R. Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484-497, 2017.

[7] A. A. Aburomman and M. B. I. Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Applied Soft Computing*, vol. 38, pp. 360-372, 2016.

[8] E. Aghaei and G. Serpen, "Ensemble classifier for misuse detection using N-gram feature vectors through operating system call traces," *International Journal of Hybrid Intelligent Systems*, pp. 1-14, 2017.

[9] W.-C. Lin, S.-W. Ke, and C.-F. Tsai, "CANN: An intrusion detection system based on combining cluster centers and nearest neighbors," *Knowledge-based systems*, vol. 78, pp. 13-21, 2015.

[10] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A new feature selection model based on ID3 and bees algorithm for intrusion detection system," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 23, pp. 615-622, 2015.

[11] M. H. Aghdam and P. Kabiri, "Feature Selection for Intrusion Detection System Using Ant Colony Optimization," *IJ Network Security*, vol. 18, pp. 420-432, 2016.

[12] N. Pandeeswari and G. Kumar, "Anomaly detection system in cloud environment using fuzzy clustering based ANN," *Mobile Networks and Applications*, vol. 21, pp. 494-505, 2016.

[13] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Systems with Applications*, vol. 42, pp. 2670-2679, 2015.

[14] F. Kuang, W. Xu, and S. Zhang, "A novel hybrid KPCA and SVM with GA model for intrusion detection," *Applied Soft Computing*, vol. 18, pp. 178-184, 2014.

[15] F. Kuang, S. Zhang, Z. Jin, and W. Xu, "A novel SVM by combining kernel principal component analysis and improved chaotic particle swarm optimization for intrusion detection," *Soft Computing*, vol. 19, pp. 1187-1199, 2015.

[16] M. A. M. Hasan, M. Nasser, S. Ahmad, and K. I. Molla, "Feature selection for intrusion detection using random forest," *Journal of information security*, vol. 7, p. 129, 2016.

[17] I. Ahmad, "Feature selection using particle swarm optimization in intrusion detection," *International Journal of Distributed Sensor Networks*, vol. 11, p. 806954, 2015.

[18] B. Aslahi-Shahri, R. Rahmani, M. Chizari, A. Maralani, M. Eslami, M. Golkar, *et al.*, "A hybrid method consisting of GA and SVM for intrusion detection system," *Neural computing and applications*, vol. 27, pp. 1669-1676, 2016.

[19] J. Kevric, S. Jukic, and A. Subasi, "An effective combining classifier approach using tree algorithms for network intrusion detection," *Neural Computing and Applications*, vol. 28, pp. 1051-1058, 2017.

[20] E. De la Hoz, E. De La Hoz, A. Ortiz, J. Ortega, and B. Prieto, "PCA filtering and probabilistic SOM for network intrusion detection," *Neurocomputing*, vol. 164, pp. 71-81, 2015.

[21] O. Al-Jarrah and A. Arafat, "Network intrusion detection system using neural network classification of attack behavior," *Journal of Advances in Information Technology Vol*, vol. 6, 2015.

[22] "http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html."

[23] "https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-IDS-Datasets/."

Authors' Profiles

Ruby Sharma is a research scholar in the department of CSE, School of Computing & I.T. in Manipal University Jaipur. Currently she is working as an associate professor in Institute of Information Technology and Management, New Delhi. She completed M.Tech in Information Technology in the year 2010. She did her B.Tech. In Electronics in the year 2001 from Aligarh Muslim University. She has more than 16 years of rich experience in industry, research and academics. She has publications in National and International journals/ conference proceedings. She is lifetime member of ISTE.

Dr. Sandeep Chaurasia is working as Associate Professor in the department of CSE, School of Computing & I.T. in Manipal University Jaipur. He completed his PhD (Engineering) in 2014 in the area of Supervised Machine Learning and M.Tech in Computer Science in the year 2009. He has done his B.E. in Computer Engineering in the year 2006 from Rajasthan University. He has more than ten years of rich experience in industry, research and academics. He has publications in National and International journals/ conference proceedings. He is also member of reviewer board of various journals and technical program committee of several reputed conferences. He is active member of IEEE, LMCSI. ACM, MIRL, UACEE etc.

How to cite this paper: Ruby Sharma, Sandeep Chaurasia,"An Integrated Perceptron Kernel Classifier for Intrusion Detection System", International Journal of Computer Network and Information Security(IJCNIS), Vol.10, No.12, pp.11-20, 2018.DOI: 10.5815/ijcnis.2018.12.02