Comparative study between (SVM) and (KNN) classifiers by using (PCA) to improve of intrusion detection system

Nafea Ali majeed Alhammadi
Department of Computer Science, Shatt Al-Arab University College, Basra, Iraq

Article Info

Article history:

Received February 15, 2022
Revised April 7, 2022
Accepted April 27, 2022

Keywords:

Intrusion detection,
Support Vector Machine,
K-Nearest Neighbor,

ABSTRACT

Intrusion Detection Systems (IDSs) are efficient applications that monitor activities of specific network or system to detect any abnormal activity and then send alarms for a defined management station. However, the current IDSs generate a high number of false alarms; False Positives (FP) and False Negatives (FN), which decreases the accuracy of distinguishing attacks from normal activities. Therefore, this thesis introduces the implementation of enhanced IDS using two classifiers: PCA-SVM and PCA-KNN. The performance of the system with using these classifiers is compared using the NSL-KDD dataset to determine the optimal classifier in terms of detection rate and the number of generated false alarms. This is performed based on dividing the dataset into training and testing sets, where the Control Chart is then applied on the training set to improve the results, where it filters the data to remove the out-bound data and keep the data in the range from Mean-3sigma to Mean+3sigma.

Six evaluation metrics; FP, FN, True Positive (TP), True Negative (TN), Detection Rate (DR) and Classification Rate (CR) are computed for both classifiers for three sets of features; with and without applying a control chart. The obtained results demonstrate that the PCA-KNN based IDS with control chart offers the best detection rate with minimum number of generated false alarms for sets F2 and F3, while the PCA-SVM based IDS with control chart offers the best detection rate with minimum number of generated false alarms for F1. The average achieved detection rate for the PCA-KNN based IDS is 98.17% with control chart and 88.7738% without control chart. On the other hand, the average achieved detection rate for the PCA-SVM based IDS is 97.62% with control chart and 96.63587% without control chart. Based on these outcomes, the application of control chart enhances the detection rate and decreases the number of false alarms for both classifiers. In addition, the PCA-KNN is the best classifier to be applied on the IDS with minimum number of false alarms and highest security and detection rate. Our proposed IDSs are implemented and tested in MATLAB 2014.

Corresponding Author:

Nafea ali majeed alhammadi
Department of Computer Science, Shatt Al-Arab University College, Basra, Iraq
Email: nafealhammadi@sa-uc.edu.iq

1. INTRODUCTION

Computers and communication are considered as essential parts of human life. The world is becoming more and more interconnected with both the internet and networking techniques. Thus, the amount of the available commercial, personal, government and military data that increases the importance of network security due to vulnerability of important data. Practically, various security tools, such as firewalls, anti-viruses and policies have been proposed to reduce threats, reach statutory compliance and address the information security problems. IDSs are software applications that used to monitor the activities of networks or systems, detect unauthorized records, activities and events, such as attacks and then respond automatically to these activities. On the other hand, these systems do not completely guarantee the security issue in networks and have some restrictions.

The rapid improvements and enhancements in internet based technologies and techniques, various application domains in both computers and communication have emerged and considered as main parts of human life. The accessibility of cheap broad band, mobile technologies, and internet connectivity raised the number of connected computers to the internet. Nowadays, the world is becoming more and more interconnected with both the internet and networking techniques. Thus, the amount of the available commercial, personal, government and military data on the networking infrastructures is being increased daily where thus in turn increases the importance of network security due to vulnerability of important data and intangible intellectual property to various types of attacks and threats. [1]

However, IDSs are advanced security tools that can be used to detect various types of attacks in networks. The main problem of these systems is their low accuracy. The current IDSs are not precise enough to offer reliable detection, where this problem resulted in a high number of generated false alarms: False Positives (FP) and False Negatives (FN). This large number of false alarms makes the process of filtering out false attacks without missing real ones a real challenge. Furthermore, it makes security administrators unable to respond correctly for risks.

Both the Support Vector Machine (SVM) and K Nearest Neighbor (KNN) are some of the best classification methods that can be used to detect FP, FN and accuracy for NSL-KDD CUP 99 dataset. The SVM is supervised learning classifier, which depends on creating a hyper-plane using support vectors to separate normal from abnormal data, while the KNN is a machine learning technique that can be utilized to discover new added data for training set. In this work, both methods will be applied to the developed IDS to determine the most efficient one.

2. THEORETICAL BACKGROUND

Various researches explored the implementation of IDSs that provide details and information concerning the features of those systems, which are in turn appropriate and applicable in the detection of various types of attacks. The implementation of those systems is based on the experiences that resulted from both the development and utilization of IDS and the analysis of various kinds of threats. [1]

The main IDS characteristics are the information that utilized in the analysis, the verification and interpretation levels of protocols and the utilized approaches in finding activities, which can signify attacks. IDSs are mainly range from simple to complicated systems based on their properties. Two simple parameters can be used to represent IDS characteristics. The first one represents the general characteristics of the system, such as the ability to determine conventional expression similarity on data, but this parameter cannot define the target of that characteristic or determine its type. The second parameter can define the target of the system characteristic to decide the validity of the system characteristics. [2].

Support Vector Machine (SVM) is an advanced machine learning technique where it outperforms many other typical machine learning techniques in the various field. The SVM is a very efficient method for classifying where it determines the most optimal separating hyper-plane among classes depending on training cases. To understand this technique, suppose that there are two linearly separable classes in a certain d-dimensional space with the use of training vectors that related to two classes; $\{x_i, y_i\}$ in which $x_i \in R^d$ signifies vectors in the d-dimensional space, while $y_i \in \{-1, +1\}$ represents a class label. The purpose is the design of a hyper-plane to simplify data in an accurate way where this hyper-plane is the one that leaves the extreme margin from both classes [3]

The main idea of SVM technique is finding the hyper-plane which has the most extreme margin towards the sample object, where the margin value and the probability to inaccurately classify a feature vector are inversely related to each other. The following equation can be used to define a hyper-plane [3].

$$(w.x) + b = U \quad (4)$$

Where w is a normal to the hyper-plane, x represents a feature vector that lies on that hyper-plane and b represents the bias in which $\frac{|b|}{\|w\|}$ represents a perpendicular distance among the origin and the middle point of the hyper-plane as shown figure 1 concerning the SVM basics.

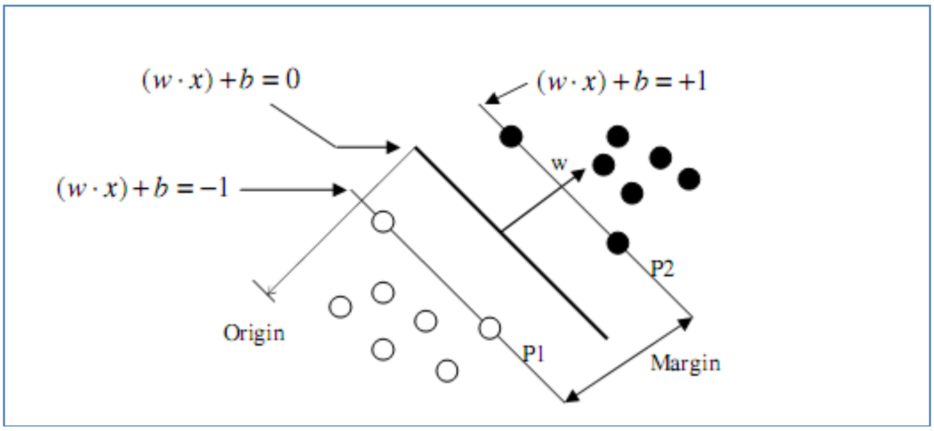

Figure 1. Shows SVM classification [5]

The purpose is to separate among two classes; open circle that stands for the class label -1 and the solid circle which stands for the +1. The lying circles on both planes; P1 and P2 represent the support vectors in which the optimal *hyper-plane* located among those two plans which are parallel to each other. The margin among those planes is $\frac{|2|}{\|w\|}$. The SVM technique should maximize the *hyper-plane* margin to get enhanced generalization. The following formulas can be used then to describe the hyper-plane of the two classes [6].

$$(\mathrm{w.x}) + \mathrm{b} = +1 \quad \mathrm{forclassy} = +1 \qquad (2)[6]$$

$$(\mathrm{w.x}) + \mathrm{b} = -1 \quad \mathrm{forclassy} = -1 \qquad (3)\ [6]$$

Practically, classes are not linearly separated. Thus, the input space must be mapped into another feature space with high dimensionality. More clearly, input vectors, such as the low-level feature vectors are mapped into a feature space H using a nonlinear conversion, $\Phi: \mathrm{R}^{\mathrm{d}} \rightarrow \mathrm{H}$.Thus, the optimal *hyper-plane* is generated in that high dimensional feature space with the use of kernel function; $\mathrm{K}(\mathrm{x}_{\mathrm{i}}, \mathrm{x}_{\mathrm{j}})$ that generated among two input vectors; $\mathrm{x}_{\mathrm{i}}$and x_{j}. This formula can be written as follows: [4]

$$\mathrm{K}(\mathrm{x}_{\mathrm{i}}, \mathrm{x}_{\mathrm{j}}) = \Phi(\mathrm{x}_{\mathrm{i}}).\Phi(\mathrm{x}_{\mathrm{j}}) \qquad (4)\ [4]$$

Polynomials kernel is one of the most common mappings, where its formula is described below: [1]

$$\mathrm{K}(\mathrm{x}_{\mathrm{i}}, \mathrm{x}_{\mathrm{j}}) = (\mathrm{x}_{\mathrm{i}}.\mathrm{x}_{\mathrm{j}} + 1)^{\mathrm{d}} \qquad (5)[1]$$

Where, d represents the polynomial degree. Another common mapping is the Radial Basis Functions (RBFs) kernel as described below: [4]

$$\mathrm{K}(\mathrm{x}_{\mathrm{i}}, \mathrm{x}_{\mathrm{j}}) = \mathrm{e}^{\frac{\|\mathrm{x}_{\mathrm{i}} - \mathrm{x}_{\mathrm{j}}\|^2}{2\sigma}} \qquad (6)[4]$$

Where,σ stands for the Gaussian sigma. As described above, the SVM technique is developed to solve binary classification problems with two class labels only; +1 and -1. This technique can be enhanced more to be used for multi-class problems. Generally, there are two approaches for SVM multi-class classification; one against all and one against one. The one against all approach includes the construction of SVMs among each class and other classes). As an example, suppose that there are four classes; C1, C2, C3 and C4, thus, four SVMs must be generated in which C1 can be classified based on classifying C1 and on C1 by the corresponding SVM and the same for other classes.

The one against one approach includes the construction of SVMs among the whole pairs of classes. As an example, suppose that there are four classes; C1, C2, C3, and C4, thus, six SVMs must be generated where those six classifiers classify [C1 or C2], [C1 or C3], [C1 or C4], [C2 or C3], [C2 or C4] and [C3 or C4].

KNN is a machine learning technique that classifies data depending on their similarity with data in the training set. This technique makes decision depending on the whole training dataset. The KNN is a simple method, which saves all obtainable cases and categorizes new data depending on a certain similarity measure. This method has been applied in various pattern recognition and statistical estimation applications. In this method, data are classified via a majority vote of its neighbors, where data are assigned to the most common class between all its K-Nearest Neighbors that measured using a certain distance function. When the number of nearest neighbors is one, then the data are assigned to that class. This method does not depend on using training data points for generalization. This means that there is no clear training phase, thus the training phase is quick. This demonstrates that this method keeps the whole training data.

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} = \frac{k[y]}{k} \quad (7)[7]$$

This database composed of 41 network connection features, where the names of those features are demonstrated in this research. The NSL-KDD dataset can be downloaded from (http://iscx.ca/NSL-KDD/). The proposed classification methods are applied on the proposed IDSs using KDD dataset. All the used features should have numeric values in order to be implemented in the classifier.

Table 1. Dataset feature

Features #	Name	Features #	Name
1	Duration	22	is_guest_login
2	protocol_type	23	Count
3	Service	24	srv_count
4	Flag	25	serror_rate
5	src_bytes	26	srv_serror_rate
6	dst_bytes	27	rerror_rate
7	Land	28	srv_error_rate
8	wrong_fragment	29	same_srv_rate
9	Urgent	30	diff_srv_rate
10	Hot	31	srv_diff_host_rate
11	num_failed_logins	32	dst_host_count
12	logged_in	33	dst_host_srv_count
13	num_compromised	34	dst_host_same_srv_rate
14	root_shell	35	dst_host_diff_srv_rate
15	su_attempted	36	dst_host_same_src_port_rate
16	num_root	37	dst_host_srv_diff_host_rate
17	num_file_creations	38	dst_host_serror_rate
18	num_shells	39	dst_host_srv_serror_rate
19	num_access_files	40	dst_host_rerror_rate
20	num_outbound_cmds	41	dst_host_srv_error_rate
21	is_hot_login		

3. THE PROPOSED IDS SYSTEM

In this work, PCA-KNN is applied and compared with PCA-SVM In the KDD dataset based IDS, both the training and testing stages are prepared through a categorization process. The resultant database includes 41 features.

NSL-KDD database is used to measure the system performance. This database includes 41 features of the network connection. In this work, the MATLAB program is used to apply system with the use of this database.

The presented IDS includes three stages; training, testing and running. In the training stage, a training database is used to train the system to recognize the normal connections from the attacked ones. Thus, the training database should have adequate information concerning connections and attacks. In this stage, a SVM classifier based feature selection method is used to recognize the most important features to be used in detecting attacks. In the testing stage, a testing database that includes connections and attacks is used in the system to measure the IDS performance, where the high performance stands for the high accuracy in determining both connections and attacks. When the system performance level is not accepted, these two stages are executed again. In the running stage, the system is used to protect the network traffic. In both the testing and running stages, the system categorizes the network traffic depending on the requested service and then depending on the chosen features.

Recently, various researches have been conducted to find optimal solutions for reducing the high dimensionality for image feature vectors. It was found that the efficient solution for reducing the high dimensionality is the application of various dimensionality reduction approaches which are classified into linear and nonlinear dimension reduction approaches. The main linear dimension reduction approaches are random projection singular value decomposition (SVD), and principal component analysis (PCA). On the other hand, the main nonlinear dimension reduction approach is the multi-dimensional scaling (MDS). Generally, the PCA approach is the most efficient and appropriate one for dimensionality reduction. It depends initially on computing both the mean vector (μ) and the covariance matrix (M) from databases using the following formulas:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (8)\ [8]$$

$$C = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T \quad (8)[8]$$

Where n represents the number of feature vectors. After computing both the mean vector and the covariance matrix, both the eigenvectors and Eigen-values of the covariance matrix M are calculated. After that, the Eigenvectors that related to the most essential Eigen-values give the principal components.

The use of PCA in this work is to reduce dimension of the data which it is 41 in KDD data, so to overcome on the huge dimension space of this data PCA is used as feature selection. On the other hand some feature has not a significant value that effect on the behavior of the packet so reduction of the dimension is considered as valuable stage especially when deal with SVM system that take long processing iterations.

A reduction process has been used to reduce the number of features in order to decrease the complexity of the IDS. It is well known that PCA has been widely used in data compression and feature selection. Feature selection refers to a process whereby a data space is transformed into a feature space, which has a reduced dimension.

In this research, PCA is applied on the training data for feature selection, then the classification is applied into normal and attack records, in training phase normal data will be taken as training data to both SVM and KNN classifiers in order to learning main features of normal data, then to filter the training data, control chart (CC) will be used as lower control chart (LCC) and upper control chart (UCC), by compute mean and standard deviation to each record. Control chart is used to control normal training data within a specific range, in order to apply testing data on same range, so it is predicted that any testing records in range of CC should be normal data and any testing records lower than LCC or upper than UCC should be attack.

On the other hand, the performance of both classifiers will be compared to each other. The expected results can be shown and compared with different practical scenarios as shown in the flow chart below.

PCA is applied on the NSL-KDD dataset to reduce dimension and feature selection in order to decrease the complexity of the IDS. It depends initially on computing the covariance matrix (C), Firstly, preprocessing to dataset from feature mapping and scaling then feature selection compute the covariance matrix (C) then data splitting to training and testing data.

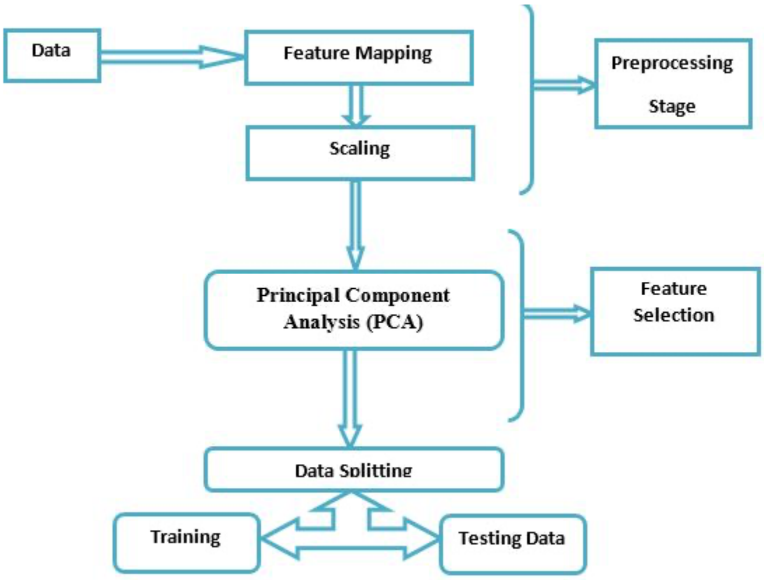

Figure 2. Flowchart of proposed IDS system based PCA

In training stage, SVM classifier based feature selection method is used to recognize the most important features to be used in detecting attacks. In the testing stage, a testing dataset that includes connections and attacks is used in the system to measure the IDS performance, where the high performance stands for the high accuracy in determining both connections and attacks then compute the FP,FN,TP,TN,DR,CR.

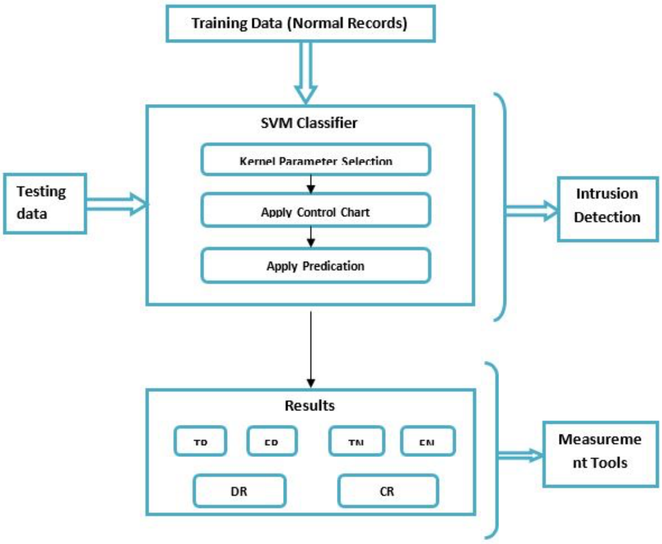

Figure 3. Flowchart of proposed IDS system based PCA-SVM

KNN is a machine learning technique that classifies data depending on their similarity with data in the training set. This technique makes decision depending on the whole training dataset. The KNN is a simple method, which saves all obtainable cases and categorizes new data depending on a certain similarity measure by using several steps .firstly, Determine k then Compute the distances among new data and the training data then Sort the distances and decide the k nearest neighbors then Collect their classes and decide the optimal class after that compute FP, FN, TP, TN, DR, CR.

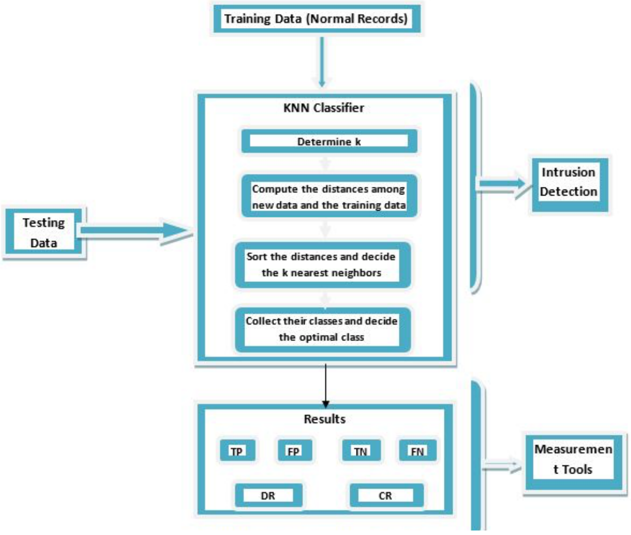
Training Data (Normal Records)
KNN Classifier
Determine k
Compute the distances among new data and the training data
Sort the distances and decide the k nearest neighbors
Collect their classes and decide the optimal class
Testing Data
Intrusion Detection
Results
TP
FP
TN
FN
DR
CR
Measuremen t Tools

Figure 4. Flowchart of proposed IDS system based PCA-KNN

The main meaning behind Support Vector Machines (SVMs) is to enable us to extract and accomplish a mixed component that maximizes the separating margin between two Confusion Matrix classes the negative and the positive. [9]

An introduction to SVM strategy founded by [34] the main goal of SVM is that it approximates the implementation of the Structure Risk Minimization (SRM) principle that in its basic structure based on statistical learning theory rather than the Empirical SRM, in the way that the classification function that SVM adopt it in the way of minimizing the Mean Square Error (MSE) all over the training data set records. Form the metrics that used in the aim to estimate the classification quality is by measuring the classification accuracy. Another important metrics that must be addressed is the need to measure the running time (computational complexity) of the intrusion detector. The computational complexity of SVMs is of linear proportion to the number of support vectors this considers as a problem in evaluating the computational complexity value since it is in a linear relation with the number of vectors.

The KNN is a simple method, which saves all obtainable cases and categorizes new data depending on a certain similarity measure. This method has been applied in various pattern recognition and statistical estimation applications. In this method, data are classified via a majority vote of its neighbors, where data are assigned to the most common class between all its K-Nearest Neighbors that measured using a certain distance function. When the number of nearest neighbors is one, then the data are assigned to that class. This method does not depend on using training data points for generalization. This means that there is no clear training phase, thus the training phase is quick. This demonstrates that this method keeps the whole training data.By taking the average of the K-neighbors nearest to the testing process, it can smooth out the impact of isolated noisy training examples.

This research in a simple view provided a methodology that provides a security solution based on K-Nearest Neighbor method and SVM. For reaching a better level in evaluation on unknown attacks, in the proposed methodology the detection of suspicious traffic using the clustering strategy well be tested integrating the SVM filter on them. Following attractive points is interesting in proposed method
1. As a first step there is a process of classifying the network traffic using SVM (support Vector Machine)
2. Then as a second step by applying, clustering based detection as a stage and prevention of intrusion on real time traffic as another stage instead of KNN.

Firstly, data from NSL-KDD are used, where the dimension and feature selection is reduced using the PCA. The data are then divided into two sets; training and testing setsThen apply the Control Chart on the training and testing dataset where both the SVM and KNN classifiers are applied on the training and testing dataset to measure the IDS performance. where the FP, FN, TP, TN, DR, and CR metrics are computed. The performance of both classifiers is compared to determine the optimal classifier that offer the lowest FP and FN.

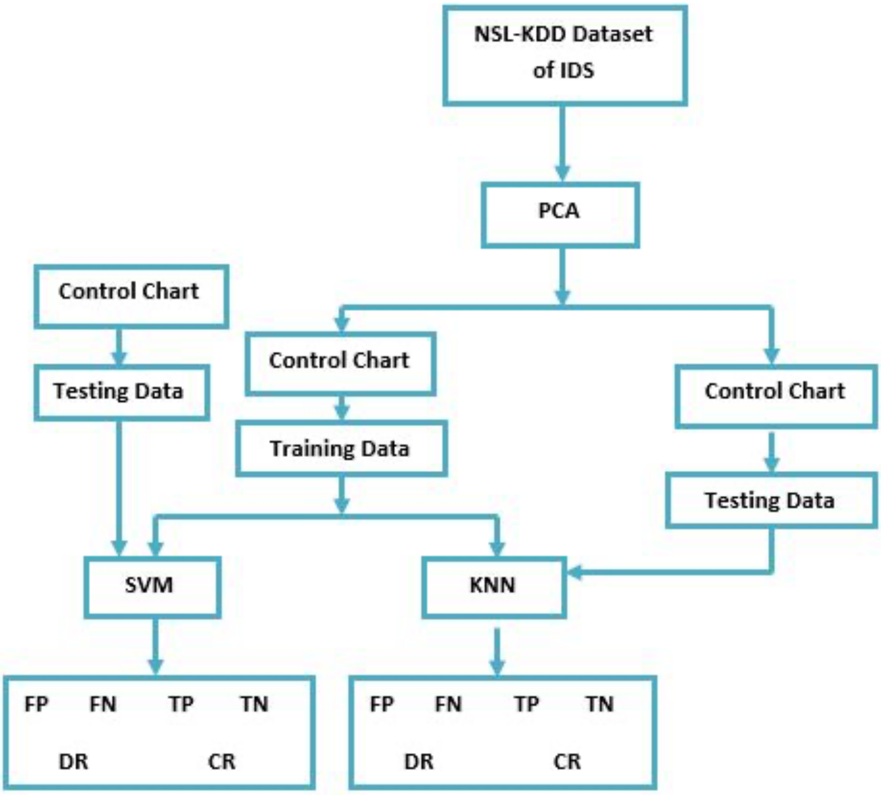

Figure 5. Flowchart of proposed IDS system based PCA-SVM and PCA-KNN

4. THE EXPERIMENT RESULTS

The testing engine is used to test the resultant training engine by using the KDD dataset and to determine if the record is an attack or not based on a specified threshold. Accuracy and results of tests depend on the databases, features and threshold value. The following percentage expressions are used in the analysis of data. [10]

- True Negative (TP): Normal records which are correctly classified,
- True Positive (TP): Attack records which are correctly classified,
- False Positive (FP): Normal records which are incorrectly classified as attacks,
- False Negative (FN): Attack records which are incorrectly classified as normal.

By using these expressions, both the detection rate and classification rate can be represented as follows:

$$Detection\ Rate\ (DR) = \frac{TP}{TP+FN} \quad (9)[10]$$

$$Classification\ Rate\ (CR) = \frac{TP+TN}{TP+TN+FPFN} \quad (10)[10]$$

The resultant most effective 11 features, which have the highest Eigen values after applying the PCA, are illustrated in the following table. Other features are considered as noisy ones.

Table 2. Best output 11 features from after applying the PCA

Features #	**Name**	**Eigen value**
4	Flag	**$1.043*10^{11}$**
5	src_bytes	**$2.168*10^{9}$**
10	Hot	**$1.0168*10^{6}$**
11	num_failed_logins	**$2.532*10^{5}$**
23	Count	**$1.1297*10^{4}$**
24	srv_count	**$6.188*10^{3}$**
29	same_srv_rate	**$5.020*10^{3}$**
31	srv_diff_host_rate	**$1.608*10^{3}$**
33	dst_host_srv_count	**$1.263*10^{3}$**
38	dst_host_serror_rate	**3.2156**
41	dst_host_srv_error_rate	**2.8067**

Thus, the first set of features includes 11 features; F1: [4,5,10,11,23,24,29,31,33, 38,41]. As shown in the table above, the two features that have the less Eigen values are 38 and 41. Therefore, these two features are removed to have a second set of features; F2: [4,5,10,11,23,24,29, 31,33]. The same process is applied then in this set where the two features that have the less Eigen values are 31 and 33. Therefore, these two features are removed to have a third set of features; 7 features F3: [4,5,10,11,23, 24,29].

Results of Applying the PCA-KNN Based IDS on F1

The obtained results of applying the PCA-KNN based IDS without control chart on F1 that includes 11 features from the NSL-KDD dataset; [4,5,10,11,23,24,29,31,33, 38,41] is shown in the following table .

Table 3. Results of applying the PCA-KNN based IDS on F1 without control chart

TN	FP	TP	FN	DR	CR
89.8545%	10.1455%	91.7786%	8.2214%	91.7786%	90.8165%

The results above demonstrate that the system has 8.2214% and 10.1455% FN and FP percentages; respectively, which stand for false alarms. Thus, the related records for these alarms should be removed from the dataset. Conversely, the system has 91.7786% and 90.8165% detection and classification rates, respectively.

Results of Applying the PCA-KNN Based IDS on F2

The following figure and table show the obtained results of applying the PCA-KNN based IDS without control chart on F2 that includes 9 features from the NSL-KDD dataset; [4,5,10,11,23,24,29,31,33].

Table 4. Results of applying the PCA-KNN based IDS on F2 without control chart

TN	FP	TP	FN	DR	CR
89.8165%	10.1835%	91.5718%	8.4282%	91.5718%	90.6942%

It can be noticed that the system has 8.4282% and 10.1835% FN and FP percentages; respectively. Conversely, the system has 91.5718% and 90.6942% detection and classification rates, respectively.

Results of Applying the PCA-KNN Based IDS on F3

This section demonstrate the obtained results of applying the PCA-KNN based IDS without control chart on F3, which includes 7 features from the NSL-KDD dataset; [4,5,10,11,23,24,29]. The following table show the achieved outcomes.

Table 5. Results of applying the PCA-KNN based IDS on F3 without control chart

TN	FP	TP	FN	DR	CR
89.3335%	10.6665%	82.9710%	17.0290%	82.9710%	86.1522%

5. CONCLUSION

This work offers an enhancement for current Intrusion Detection Systems (IDSs), which suffer from high numbers of generated false alarms that are represented by high percentages of both False Positives (FP) and False Negatives (FN). This is performed based on applying both classifiers; Support Vector Machine (SVM) and K Nearest Neighbor (KNN) on an IDS using the MATLAB program to determine the best classifier that decrease the number of generated false alarms, enhance the network security and improve the detection rate of various types of attacks.

For further enhancements, the Principal Component Analysis (PCA) technique is combined with both classifiers to offer two enhanced classifiers; PCA-SVM and PCA-KNN. The NSL-KDD dataset is used to evaluate and measure the system performance after applying both classifiers based on dividing it into two sets; training and testing. The use of the PCA technique offers an enhancement for these two sets based on reducing their dimensionalities and selecting the optimal features.

REFERENCES

[1] M. Dacier, & D. Alessandri," A Vulnerability Database", Computer Security Journal. (1999)

[2] G. R.Lanckriet, M.Deng, , N.Cristianini, , M. I.Jordan, , & W. S. Noble," Kernel-based data fusion and its application to protein function prediction in yeast",In Pacific symposium on biocomputing. 300-311.(2004)

[3] R. Lippmann, D. Fried, and I. Graf,"Evaluating Intrusion Detection Systems: the '1998 DARPA off-line Intrusion Detection Evaluation", In: Information Survivability Conference and Exposition, IEEE, pp 12-26.(2000)

[4] D .Alessandria. "Attack-Class-Based Analysis of Intrusion Detection Systems.University of Newcastle upon Tyne School of Computing Science" .(2004)

[5] Y. Bhavsar, W.B, & C. Kalyani ," Intrusion Detection System Using Data Mining Technique: Support Vector Machine", International Journal of Emerging Technology and Advanced Engineering. (2013)

[6] R. Pedersen, & M. Schoeberl,"An embedded support vector machine. In Intelligent Solutions in Embedded Systems", IEEE . 1-11 .(2006)

[7] G.Dewaele , K. Fukuda & P. Borgnat,. "Extracting hidden anomalies using sketch and non Gaussian multi-resolution statistical detection procedures".(2007)

[8] J. Hu," Host-Based Anomaly Intrusion Detection'', Springer. 235-255.(2010)

[9] A. Hofmeyr, S. Forrest & A. Somayaji," Intrusion detection using sequences of system calls",Journal of Computer Security. 6: 151–180.(1998)

[10] Y. Ming," Real Time Anomaly Detection Systems for Denial of Service Attacks by Weighted k-Nearest Neighbor Classifiers", Expert Systems with Applications. (2011)

[11] V. Chandola, & V. Kumar, "Anomaly Detection: A Survey, ACM Computing Surveys", 1-72, (2009)

[12] R. Chen, K. Cheng, Y. Chen, and C. Hsieh," Using rough set and support vector machine for network intrusion detection", International Journal of Network Security & Its Applications (IJNSA). (2009)

[13] D. Faria, "Scalable location-based security in wireless networks khan.Federal Information Processing Standards Publication 191 (FIPS PUB 191). (1994)", Guideline for the Analysis Local Area Network Security.(2006)

[14] T. S. Furey , N.Cristianini, , N. Duffy, , D. W.Bednarski, , M.Schummer, , & D. Haussler," Support vector machine classification and validation of cancer tissue samples using microarray expression data",Bioinformatics.906-914.(2000)

[15] N. I. Ghali, "Feature Selection for Effective Anomaly-Based Intrusion Detection". IJCSNS International Journal of Computer Science and Network Security. 9 (3) .(2009)

[16] P. Gogoi, D.K.Bhattacharyya, B. Borah & J. k. Kalita," MLH-IDS: A Multi-Level Hybrid Intrusion Detection Method",The Computer Journal Advance. (2013)

[17] Q. P. He, , & J. Wang,." Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes. Semiconductor manufacturing". IEEE.345-354. (2007)

[18] P. T. Htun, and K. T.Khaing," Anomaly Intrusion Detection System using Random Forests and k-Nearest Neighbor", International journal ssrg. (2015)

[19] A. Lakhina, M. Crovella, & C. Diot,"Mining anomalies using traffic feature distributions". 21–26.(2005)

[20] J. Leung," Vulnerability Management – A Guide to Managing Internal and External Threats".(2008)

[21] W.Li, , P.Yi, , Y. Wu, , L.Pan, , and J.Li, "A New Intrusion Detection System Based on KNN Classification Algorithm in Wireless Sensor Network", Hindawi Publishing Corporation Journal of Electrical and Computer Engineering. (2014).

[22] Y. Li, and L. Guo,"An active learning based TCM-KNN algorithm for supervised network intrusion detection". Computers & security. 459-467.(2007)

[23] Y.Li, , B. Fang, L .Guo, & Y. Chen,"Network Anomaly Detection Based on TCM-KNN Algorithm",13-19.(2007)

[24] W. K.Liu, S.Jun, , & Y. F. Zhang," Reproducing kernel particle methods". International journal for

numerical methods in fluids.1081-1106.(1995)

[25] S.Mukkamala, G. Janoski, , and A. Sung,”Intrusion detection using neural networks and support vector machines. In Neural Networks”, IEEE. International Joint Conference on. 1702-1707.(2002)

[26] S. Mulay, P. Devale, G. Garje,”Intrusion detection system using support vector machine and decision tree”, International Journal of Computer Applications. (2010)

[27] P.Porras, D. Schnacenberg, S. S. Chen, & F. Wu, ”The Common Intrusion Detection Framework Architecture”, Journal of Computer Security. (2000)

[28] Shailendra and Sanjay,” An ensemble approach for feature selection of Cyber Attack Dataset”, International Journal of Computer Science and Information Security. (2009)

[29] G. Tandon, & P.K. Chan,” In the Florida Artificial Intelligence Research Society Conference”. 405-410.(2005)

[30] J.Wang, , X.Hong, , R. R. Ren, & T. H. Li,” A Real-time Intrusion Detection System Based on PSO-SVM”,Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009). 319- 321.(2009)

[31] J. Xu & C.R. Shelton,” In European Conference on Machine Learning”.(2008)

[32] J.Yao, S.Zhao, , & L. Fan,” An Enhanced Support Vector Machine Model” .(2015)

[33] N.Ye, S.M. Emran, Q. Chen, & S. Vilbert, “Multivariate statistical analysis of audit trails for host-based intrusion detection”, Transactions of Computers, 51 (7): 810–820. (2002)

[34] V. N. Vapnik,” Statistical Learning Theory, John”, Inc.(1998)

[35] Y. Liao and V. R. Vemuri. “Use of K-Nearest Neighbor Classifier for Intrusion Detection” , Computers and Security, pp 439-448.(2002)

[36] H.N.K.AL-Behadili,”Classification algorithms for determining handwritten digit”, Iraq J. Electrical and Electronic Engineering, Vol.12 No.1 , 2016.

[37] H.N.K.AL-Behadili, K.R.Ku-Mahamud and R.Sagban, ”Hybrid Ant Colony Optimization and Genetic Algorithm for Rule Induction”, J. of Computer Science, Volume 16 No. 7, 2020, 1019-1028.