Received: October 16, 2017

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Enhancing the Performance of an Intrusion Detection System Through Multi-Linear Dimensionality Reduction and Multi-Class SVM

Bukka Narendra Kumar[1*] Mantena S. V. Sivarama Bhadri Raju[2] Bulusu Vishnu Vardhan[3]

[1]*Department of Computer Science Engineering,*
Sri Sai Jyothi Engineering College, Hyderabad, Telangana, India
[2]*Department of Computer Science Engineering,*
Sagi Ramakrishnam Raju Engineering College, Bhimavaram, Andhra Pradesh, India
[3]*Department of Computer Science Engineering,*
Jawaharlal Nehru Technological University College of Engineering Manthani, Peddapalli, Telangana, India
* Corresponding author's Email: bnkphd@gmail.com

Abstract: With the huge development of the usage of computer over network and advancement in applications running on various platforms captures the attention towards network security. The Intrusion Detection System (IDS) plays a vital role in detecting anomalies and attacks in the network. Earlier approaches of IDS relied on Machine Learning (M L) techniques. Due to some limitations, a better approach is needed. A combination of Machine Learning techniques and data preprocessing is an effective approach for IDS. In this work, a new dimensionality reduction technique combined with the Multi-class SVM (Support Vector Machine) is proposed for intrusion detection. In the proposed model, Multi-Linear Dimensionality Reduction (ML-DR) is proposed as a feature extraction technique to reduce the dimension in order to shorten the training time. A Multi-class SVM (M-SVM) is used to detect whether the action is an attack or not. Here the Multi-class SVM is adopted to perform multi-attack classification in a layered fashion. Radial Basis Function Kernel is used as a SVM kernel. NSL-KDD data set is used for the performance evaluation of the proposed approach. The performance metrics such as classification accuracy, false alarm rate and the correlation coefficient are evaluated to measure its efficiency. In comparison with other detection approaches, the experimental results show that the proposed model outperforms the higher classification accuracy.

Keywords: Intrusion detection system, Multi-linear dimensionality reduction, Support vector machine, Particle swarm optimization, accuracy, correlation coefficient, NSL-KDD.

1. Introduction

The rapid growth in technology leads to a possibility where computers and networks are under threat from worms, viruses and attacks. The use of devices connected to Internet is increasing every year very rapidly. The increase in the number of network devices has led to an increase in unauthorized activity, not only from external attackers, but also from internal attackers, through various types of intrusions. An intrusion is defined as a set of actions that compromise the integrity, confidentiality or availability of a resource, and is a type of attack that attempts to bypass the security mechanism of a computer system. Intrusion detection [1] is the process of monitoring and analyzing the events occurring in a computer system in order to detect signs of security problems. Mainly there are two types of IDSs, Misuse Detection and Anomaly Detection. The Misuse Detection [2] is based on the model already defined, and thus has low False Positive Rate (FPR), since it is already possess the knowledge about the attacks. One of the main problems with Misuse Detection is that, it is not able to detect the attacks which are not defined in the model. Hence the attacks that are not known are unnoticed by the security system and produce a

false negative. In Anomaly Detection [3], the normal behavior is modeled and thus the patterns deviating from normal behavior were detected easily. Anomaly Detection detects unknown attacks, because its main aim is to model good behavior. Anomaly Detection is based on the number of legitimate behaviors in the network, which deviates from the number of anomalous behavior. In that way, the one which is out of normality model is termed as an attack. This is also a problem because every normal behavior that is deviating from the modeled normal behavior was treated as anomalous, and thus causes more false positives. One important thing to notice is that the normal behavior changes daily and depends on the environment. Thus, generating a behavioral model from multidimensional data is a complex issue.

Data preprocessing and detection are the two phases involved in the intrusion detection system. Data Preprocessing converts the network traffic into the sequences of observations where every observation is represented as a feature vector. Further, the obtained features are processed for detection to detect whether it is an attack or not. Since the anomaly based IDS observes the characteristics of network traffic, the features of network traffic are in such a way that the overall FPR found to be low. Though there are so many data preprocessing approaches, every approach has its own drawbacks.

This paper proposes a new data preprocessing approach by which the system gets sufficient and more useful information to achieve the objective of low FPR and more accuracy. The proposed data preprocessing method transforms a higher dimensional feature space into lower dimensional feature space with less information loss. Here the projection matrix derived is to transform the optimal and covers the entire information of original dataset within lesser space. Further, the proposed system adopts a multi-class classification unit to an increase accuracy. Rest of the paper is organized as follows: Section 2 illustrates the details of earlier approaches. The complete detail of proposed model is illustrated in section 3. Section 4 describes the details of experimental results and finally the conclusions are provided in section 5.

2. Literature survey

Many intrusion detection models were proposed earlier to overcome the restrictions of anomaly detection model. This section analyzes traditional intrusion detection approaches. Due to the large dimensionality of network traffic, many intrusion detection models were developed with feature selection as a data preprocessing step. Sung et al. [4] developed an integrated feature selection model by combining SVM and Artificial Neural Network. For every instance of experiment, one feature is eliminated; thereby a reduced feature set is processed by which the system achieved an enhanced performance. Finally, sung processed only 34 features instead of 41 features and also achieved a significant performance enhancement in the performance of intrusion detection. Zaman et al. [5] developed a lightweight IDS based on a new feature selection technique. Zaman employed a fuzzy enhanced support vector decision [16] to obtain an enhanced performance. Amiri et al. [6] proposed a new feature selection approach according to the mutual information and the linear correlations between the features. The proposed approach resulted in a better accuracy particularly for minority attacks like U2R and R2L. Senthilnayaki et al. [7] developed an IDS model through a new feature selection technique based on the Information Gain ratio. This model combined two classification techniques such as Rule Based Classification and SVM to identify the label of class. This approach achieved a better accuracy levels only for DOS attacks.

Saxena et al. [8] considered the information gain ratio as a main aspect for feature selection criterion and developed an SVM integrated IDS with particle swarm optimization (PSO) [17] as an optimization technique. Here the SVM is accomplished for label classification. Though this approach achieved better result, the computation complexity over the deployment of SVM with PSO is not analyzed which is a crucial factor to explore the performance. Farrahi et al. [9] developed an IDS based on the k-means clustering and various machine learning algorithms such as Naïve Bayes classifier, SVM and OneR algorithms are used as classifiers. Though this approach achieved a greater accuracy for DOS attacks, the false alarm rate is higher for U2R and R2L attacks.

Along with the feature extraction techniques some researches focused to reduce the dimensionality of feature set without any information loss. Liu. G [10] built anomaly detection model which performed feature selection based on Principal Component Analysis (PCA) and classification by Neural Networks. This approach extracts only 22 features from the 41 features. Principal Components are the features those possess higher Eigen values. The main advantage of this approach is the reduced computational time due to

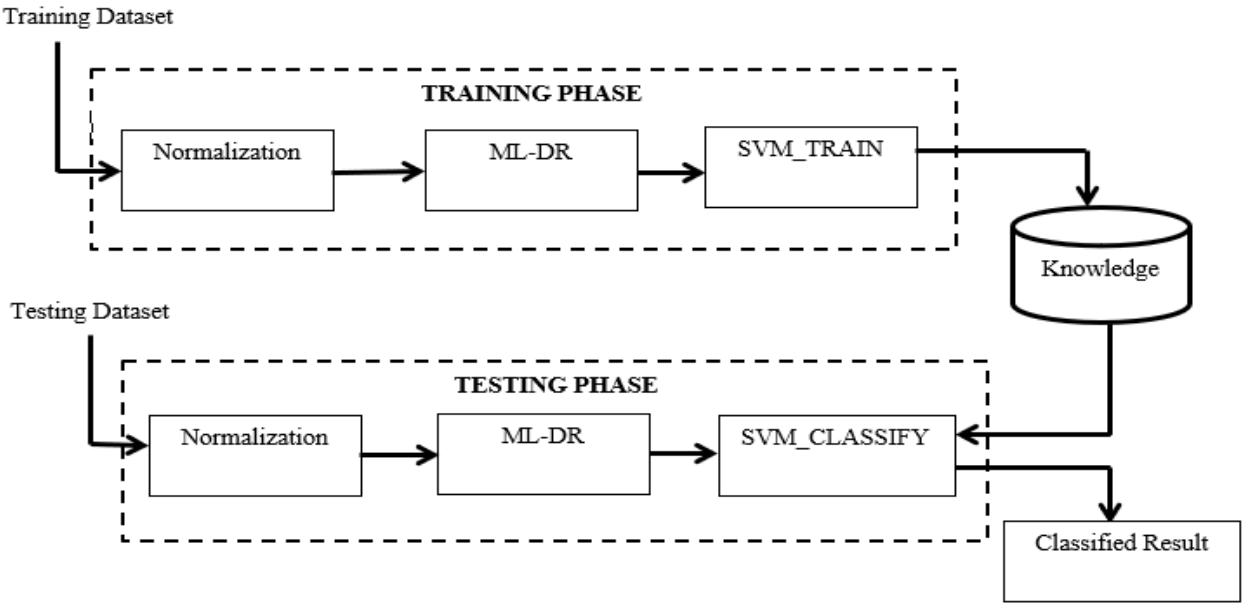

Figure.1 Block diagram of proposed approach

the reduced feature count. However, the selection of principal components is not globally optimal as a certain set of features are only processed. Kuang [11] combined the Kernel Principal Component Analysis (KPCA) and SVM for intrusion detection. Multi-layer classifier model is accomplished to determine the attacks.

Sumaiya et al. [12-14] investigated various feature selection approaches along with SVM to construct a hybrid IDS models to achieve the reduced false alarm rate. Kasliwal et al. [15] proposed a hybrid model by combining the Latent Dirichlet al.location (LDA) with Genetic Algorithm (GA) [16]. Here the LDA performs the identification of an optimal set of attributes for classification and the GA is used for optimization through the mutation and cross over operations of initial population. Sumiya Thaseen Ikram et al. [13] proposed a new feature selection approach based on the chi-square and multi class SVM. However, the main drawback with chi-square based feature selection is it does not provide much information about the relationship between the features. If the relation between the features of different intrusions is not known, entire dataset needs to be analyzed by which the processing time will be increased.

3. Proposed model and methodology

To overcome the drawbacks of conventional approaches, this work proposed a new anomaly detection model, which accomplishes in two phases. The first phase involves data preprocessing and the second phase involves the classification through Multi-Class SVM classifier. Fig. 1 represents the model of the proposed work it consists of two phases Training and Testing.

3.1 Data preprocessing

In Data Preprocessing, to reduce computational time and better classification results, ML-DR technique is used. PCA, one of the most popular dimension reduction techniques, was used in many applications like signal processing, image processing, bankruptcy and market analysis problems. Although PCA extracts features that are the most efficient for representation, it is not useful for discrimination. ML-DR selects an optimal projection matrix through which the higher dimensional feature space is projected into lower dimensional feature space, while preserving significant information for data classification. Initially, the data conversion is carried out to formulate the entire data set into a unique format. Further, the proposed ML-DR technique is applied over the data set to reduce the dimensions.

3.1.1. Data conversion

Let’s consider a data set X with N number of instances and for every instance is formulated through P number of independent features. In this paper, the X is assumed to be a NSL_KDD data set and the total features at every instance are 41. In this NSL_KDD data set, all the features don’t have a unique data format. To process the data set, all the features need to be preprocessed and has to convert

Table 1. Details of symbolic features of NSL_KDD dataset

Protocol type	Flag	Service
TCP, UDP, ICMP, ARP	OTH, REJ, RSTO, RSTOS0, RSTR, RSTRH, SHR, SF S0, S1, S2, S3, SH	Aol, http_443, http_8001, http_2784, domain_u, ftp_data, auth, bgp, courier, tftp_u, uucp_path, csnet_ns, ctf, daytime, time, discard, domain, echo, eco_i, ecr_i, efs', exec, finger, gopher, harvest, hostnames, http, imap4, IRC, iso_tsap, klogin, kshell, ldap, link, login, smtp, mtp, name, netbios_dgm, netbios_ns, netbios_ssn, netstat, nnsp, nntp, ntp_u, other, pm_dump, pop_2, pop_3, printer, private, red_i, remote_job, rje, shell, sql_net, ssh, sunrpc, supdup, systat, telnet, tim_i, urh_i, urp_i, uucp,ftp, vmnet, whois, X11,Z39_50.
Total = 4	13	70

into a unique format. Among the available 41 features of NSL_KDD dataset, only three features (Protocol type, Service and flag) are 'symbolic' in nature and remaining features are 'continuous' in nature. Table.1. Gives the details of total protocol types, flags and services of NSL_KDD dataset.

Hence these three features are processed for conversion form symbolic to continuous through the following procedure.

Step 1: Feature = **{'Protocol', 'Flag', 'Service'}**

Step 2: Choose Features one by one

Step 3: Measure the length of Features selected, L.

Step 4: Perform String comparison between the features of dataset (*X*) with the Features of step 2.

Step 5: Add all the results obtained in the step 4. The resultant formula for step 4 and step 5 is given as

$$M_i = \sum_{i=1}^{L} strcmp(Feature(i), Feature\_X)$$

Where $Feature\_X$ denotes the specified feature's field in the dataset *X*, *L* is the length of feature set (Protocol type, flag, service) specified in table.1.

Step 5: evaluate the Probabilities of Feature Type, PFT as

$$PFT_i = \frac{M_i}{Length(X)}$$

Step 6: end the process

3.1.2. Dimensionality reduction

In the proposed ML-DR technique, initially the data is grouped into P number of groups, where P is the number of possible attack types which need to be detected. Assume a data set *X* with *N* instances as $X = \{x_1, x_2, x_3, \dots, x_N\}$. After defining the number of attack types (for example, normal, DOS, Probe, U2R, R2L) *P*, select distinct P instances from the dataset *X*. Further, a similarity deviation is performed over the entire data set *X* with the selected P instances. Based on the similarity deviation between the selected random instances and the original instances, all the instances are grouped into *P* groups $(G_1, G_2, \dots, G_P)$. In the proposed ML-DR model, two distributed matrices are defined, the first one D_B which defines the distribution of instances between groups and the second one D_W which defines the distribution of instances within the group. The mathematical formula of the matrix D_B is represented as

$$D_B = \sum_{G=1}^{P} (\chi_G - \bar{X})(\chi_G - \bar{X})^T \quad (1)$$

Where $\bar{X}$ is the mean of the entire data set and is defined as,

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} x_i \quad (2)$$

And χ_G is the sample mean for group G_i, defined by

$$\chi_G = \frac{1}{N_P} \sum_{j=1}^{N_P} x_j \quad (3)$$

Where the term *N* in (2) is the total number of instances and the term N_P in (3) is the total number of samples in the P^{th} group.

Similarly, the matrix which defines the distributions between the instances within the group, D_w, defined as

$$D_w = \sum_{G=1}^{P} \sum_{i=1}^{N_P} (x_{i,G} - \chi_G)(x_{i,G} - \chi_G)^T \quad (4)$$

Then the obtained distribution matrix D_w is subjected to dimensionality reduction to remove the unnecessary feature information while preserving the significant information. Thus the optimal discriminating vector ϕ through the maximization of the following criterion;

$$J(\phi) = \frac{\phi^T D_B \phi}{\phi^T D_w \phi} \quad (5)$$

Thus, by maximizing the Eq. (5) we can achieve only one discriminating vector ϕ, which can also be denoted by ${D_w}^{-1}$ due to the rank limitation of D_B.

On the other hand, there is a possibility of non-inverse for D_w by which the singularity problem arises. Thus, the rank limitation problem does not able to give much discriminant vectors by hindering the information which makes the classification results not much effective. Further, the singularity problem needs a regularization task to remove it. Here a modified version of discriminant analysis is proposed to overcome the above shortcomings. The modified Dimensionality reduction technique is described as follows:

$$J(\phi) = \frac{tr(\phi^T \widetilde{D}_B \phi)}{tr(\phi^T \widetilde{D}_w \phi)} \quad (6)$$

Where

$$\widetilde{D}_B = \sum_{r=1}^{c} \frac{J_r}{N} \sum_{i=1}^{J_r} \sum_{\substack{j=1, \\ j \notin I_r}}^{N} (x_{ri} - \chi_j)(x_{ri} - \chi_j)^T \quad (7)$$

A new *c*-class between-class scatter matrix and

$$\widetilde{D}_w = \sum_{r=1}^{c} \sum_{j=1}^{N_r} \sum_{p=1}^{N_r} (x_{rj} - \chi_p)(x_{rj} - \chi_p)^T + \rho \sum_{r=1}^{c} \left(\frac{N_r}{N}\right)^2 \sum_{j=1}^{N} \sum_{p=1}^{N} (x_j - \chi_p)(x_j - \chi_p)^T \quad (8)$$

A new *c*-class within class scatter matrix. $N = \sum_{r=1}^{c} N_r$ denotes the total number of features, N_r the number of features in the r^{th} class in c_r. x_{ri} is the ith feature in the c_r and x_j is the jth features in the *C*, $C = \cup_{r=1}^{c} c_r$. $I_r = \{k | x_j \in c_r\}$ is an index set of c_r. ρ is an arbitrary non-negative coefficient which can control the tradeoff between the two terms in $\widetilde{D}_w$ and plays a role in the relaxation of singularity of $\widetilde{D}_w$. $\Phi_d = \{\phi_1, \phi_2, \ldots, \phi_d\}$ are the optimal discriminating vectors obtained through the maximization of Eq. (6). After solving the feature extraction optimization issue, the classification is carried out easily over the low dimensional feature space by projecting the original feature space on to the optimal projection matrix.

3.2 Classification

Since there exist multiple types of attacks in the network, Multi-class classification is required. SVM is one of the most popular non-linear classifier which gives optimal performance in the IDS. At a time SVM classifies only two classes. When the SVM is applied on the multiple way, the multiple types of attacks were detected. Binary tree, one-against-all and one-against-one are [18-20, 24-25] the most popular techniques in SVM multiclass classification. The one-against-one SVM classification technique requires *k(k-1)/2* two-class SVM classifiers where everyone is trained on data for two classes. The one-against-all SVM classification technique requires '*k*', two-class SVM classifiers to perform the detection. Finally, the binary tree SVM technique requires *k-1*, two class SVM classifiers for a test of *k* classes. Thus, the binary tree SVM classification technique is used here to perform intrusion detection. Based on the properties of various intrusion detection types, four SVM classifiers are implemented here to detect five states such as normal State and four intrusion states (DOS, Probe, U2R and R2L). Initially the first SVM (SVM1), SVM1 is trained to classify the total dataset into normal state and intrusion state.

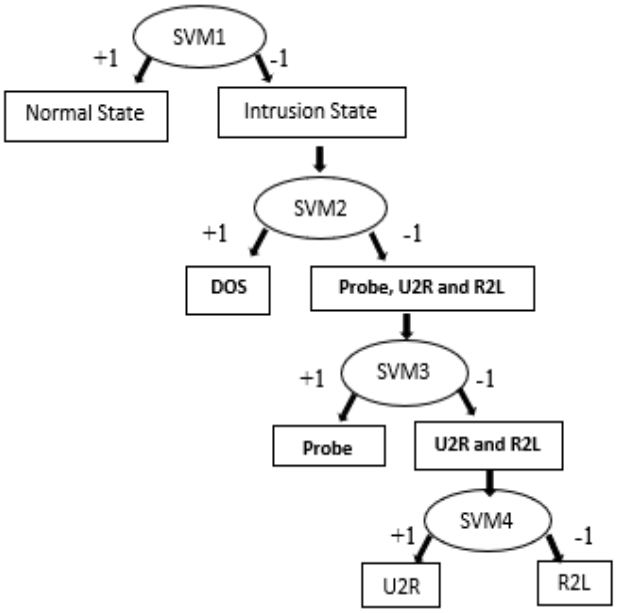

Figure.2 The scheme of intrusion detection based on Multi Class SVM Model

During this classification, the SVM produces only +1 and -1 as the output. At SVM1, if the sample representing the normal state, the output of SVM1 is +1 otherwise it is -1. Further, the obtained samples of intrusion state are given as input to SVM2. SVM2 is trained to classify the DOS attack from the intrusion state. When the input sample of SVM2 is representing the DOS, the output of SVM2 is +1; otherwise -1. Next the third SVM (SVM3), SVM3 is trained to classify the Probe from the U2R and R2L. When the input sample of SVM3 is representing the Probe, the output of SVM3 is +1; otherwise -1. Finally, the fourth SVM (SVM4), SVM4 is trained to classify the U2R from R2L. When the input sample of SVM4 is representing the U2R, the output of SVM4 is +1; otherwise -1. In

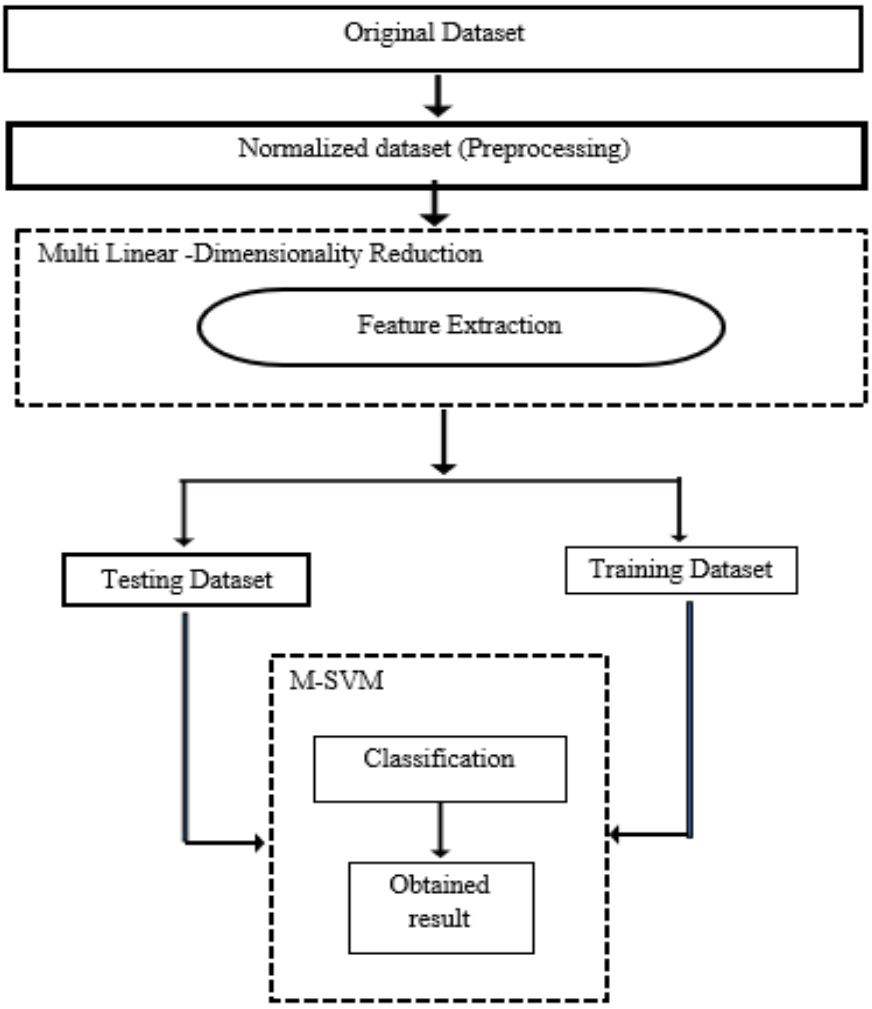

Figure.3 Procedure of the proposed model

this way, the multi-class SVM is achieved. Here all the SVMs are accomplished using Radial Basis Function (RBF) kernel. The basic principle of intrusion detection model based on Multi-Class SVM is shown in Fig. 2.

After extracting the features, the SVM is accomplished for training of the dataset. At the optimal solution, decision function using SVM is modeled as

$$f(t) = sgn\left(\sum_{i=1}^{p}(\alpha_i - \hat{\alpha}_i)K(t_i, t_j) + b\right) \quad (9)$$

Where $\alpha_i \; and \; \hat{\alpha}_i$ are the Lagrange multiplier coefficient for the i^{th} sample, $K(t_i, t_j)$ is the kernel function and b is an arbitrary constant.

Here the proposed approach accomplished the most popular RBF kernel as a kernel function to perform the decision making. Since the SVM with RBF kernel is declared as an effective combination from earlier studies, the RBF kernel is used as a kernel function. The mathematical formulation of RBF kernel is given as

$$K(t_i, t_j) = exp\left(\frac{-\|t_i - t_j\|^2}{\sigma^2}\right), \sigma \in R \quad (10)$$

According to the functional theory, as long as the function $K(t_i, t_j)$ satisfies Mercer's condition, it can be denoted as a positive definite kernel.

Fig. 3 shows the procedures of the proposed model for intrusion detection.

4. Experimental results

In this section, initially a complete analysis of the applied dataset is illustrated, then the IDS performance metrics were discussed and finally the performance evaluation of proposed approach is demonstrated.

4.1 NSL-KDD dataset

To evaluate the developed model, NSL-KDD benchmark dataset is used. NSL-KDD [21] dataset

Table 2. NSL-KDD data set classes distribution

Dataset	Normal	DoS	Probe	U2R	R2L	Total
Training set	10342	6223	1479	102	749	18895
Testing set	9711	7458	4421	167	1094	22854

Table 3. Confusion matrix

	Normal	Attack
Normal	TP	FN
Attack	FP	TN

is new version of KDD99 dataset and each NSL-KDD record consists of a host-to-host connection which has 41 distinguished features (e.g., protocol type, service and flag) and is labeled as normal, anomaly or one of the specific attack names. All attacks fall into four major group: DoS, Probe, U2R and R2L.In order to evaluate the effectiveness of the proposed approach, the samples are selected from the subset of NSL-KDD to form the training and testing set. There were five datasets in Table 2.

4.2 Performance metrics

In this paper, the detection rate(*DR*), false alarm rate (*FAR*) and Classification Accuracy (CA)are considered which are mostly used in literature to estimate the performance of intrusion detection. They were determined from the confusion matrix, as given in Table 3. The values obtained in Table 3 are as follows:

True Positives (*TP*): Total instances of anomalies correctly classified as anomalies

True Negatives (*TN*): Total number of normal instances correctly classified as normal

False Positives (*FP*): Total number of normal instances falsely classified as anomalies

False Negative (*FN*): Total number of anomalies wrongly classified as normal instance.

The derived metrics obtained from the confusion matrix are as follows:

$$\text{False Alarm Rate (FAR)} = \frac{FP}{TN+FP} \tag{11}$$

$$\text{Classification Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

These parameters are essential in evaluating the performance of the intrusion detection model. In addition, we consider another indicator *cc*, which denotes the correlation between the forecast result and the actual situation. It ranges from −1 to 1, where 1 implies the forecast result is fully consistent with the actual situation and −1 is on behalf of a random prediction.

$$Correlation\ Coefficinet\ (CC) = \frac{TP*TN-FP*FN}{\sqrt{(TP+FN)(TP+FP)(TN+FP)(TN+FN)}} \tag{13}$$

4.3 Results

The experimental evaluation is carried out in the MATLAB2012a environment, which was running on a Personal Computer (PC) with minimum of 8 GB RAM and 1 TB Hard Disk. The proposed approach was trained by the training set and then evaluated by given test set. After extracting the sufficient feature information through the proposed feature extraction technique, they were trained through the Multiclass SVM model. Here initially the SVM1 is trained through the entire feature set obtained at data preprocessing stage. The SVM1 separates the entire dataset into normal and intrusion sets. Then the SVM2 separates the DoS intrusion set from the intrusion set. Further the separation of Probe, U2R and R2L is carried out through SVM3 and SVM4 respectively. Here the number of attacks which need to be classified is five, where the number of required SVM classifiers is four. This reduces the computational complexity at classification unit. For the given test data, the obtained confusion matrix is shown in Table 4.

In the confusion matrix of Table 4, the diagonal value specifies the classified results for a given total test of respective specification. The classification accuracy is obtained by dividing the total classified results to the total number of samples given for testing. Further, the obtained classification accuracy, false alarm rate and correlation coefficient results for all types of attacks are illustrated below.

Fig. 4 shows the classification accuracy details of the proposed and conventional approaches. The classification accuracy of the proposed approach is observed to be higher compared to the conventional approaches for all cases of attacks. Thus, the proposed approach detects even the minority attacks such as U2R and R2L efficiently compared to conventional approaches. Compared with the classification accuracy of normal, DoS and Probe attacks, the classification accuracy of U2R and R2L is observed as low. Since the U2R and R2L are the minority attacks about which the system won't have much information. Hence for any IDS system, the accuracy obtains for the detection of minority attacks will be less compared to remaining attacks.

Table 4. Confusion matrix for the given test data set

	Normal	DoS	Probe	U2R	R2L	Total
Normal	**9297**	203	102	22	87	9711
Dos	150	**7153**	85	15	55	7458
Probe	99	67	**4199**	9	47	4421
U2R	14	10	6	**133**	4	167
R2L	113	46	71	4	**860**	1094
Total	9673	7479	4463	183	1053	

Table 5. Classification Accuracy comparison for all Attack types

	Normal	DoS	Probe	U2R	R2L
Chi-SVM [13]	95.236	94.125	94.558	74.238	76.664
SVM-GA [11]	95.446	94.853	94.472	76.213	76.928
Proposed ML-DR	95.738	95.996	94.971	79.775	78.669

Table.6 Performance Metrics Evaluation

Approach	Classification Accuracy	False Alarm Rate	Correlation Coefficient
Chi-SVM [13]	98.02	0.852	0.912
SVM-GA [11]	95.18	1.025	0.824
Proposed ML-DR	98.44	0.112	0.968

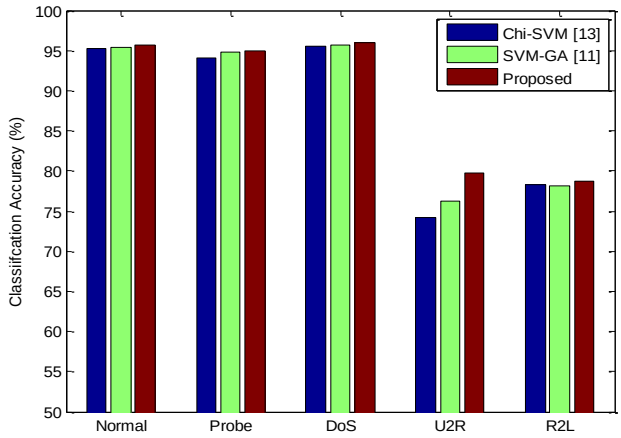

Figure.4 Classification accuracy for all types of attacks

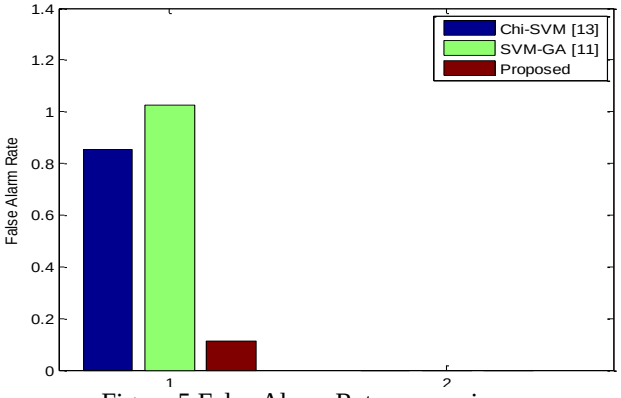

Figure.5 False Alarm Rate comparison

Similarly the proposed approach also obtained a reduced accuracy for minority attacks but it was high when compared with the accuracy of conventional approaches. This increment is due to the evaluation of multi-linear relationships between the testing and training data. Further the proposed ML-DR approach also resolves the rank limitation and singularity problem of conventional LDA, there is a possibility of an inverse for any matrix. Thus the obtained discrimination factor through the ML-DR gives more discrimination between the features set, even though the covariances are almost equal.

Further the performance metrics compared with conventional approaches are represented in Table 5. The minority attacks namely U2R and R2L are detected with good classification accuracy 79.775 and 78.669 respectively. The Comparison of the improvement in the detection rates of various attacks is as Normal: 0.34%, DoS: 1.94%, Probe: 0.47%, U2R: 6.9%, R2L: 2.57%. This shows the detection rate of proposed model on attacks is better than the conventional approaches. The details of performance metrics such as Classification Accuracy, False Alarm Rate, and Correlation Coefficent is illustrated in Table.6. It shows that all the performance metrics obtained through the proposed approach are optimal compared to the conventional approaches.

Form Table 6, it can be obsreved that the classification accuracy of proposed ML-DR is observed to be high compared to the conventional approaches. This is because, the proposed ML-DR can discriminate the differences between different types of attacks more clearly by which the system can detect the attacks more accurately. As the conventional approach Chi-SVM [13] considered chi-square distribution by which the classification of more than two classes will become more complex and is not able to achieve better accuracy. And also the conventional approaches are quite complicated to get a right –difficult formula.

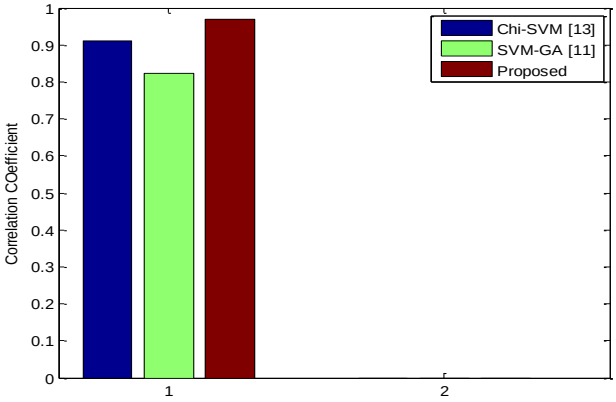

Figure.6 Correlation Coefficient Comparison

Figs. 5 and 6 show the comparative analysis of the proposed approach with the conventional approaches with respect to False Alarm Rate and Correlation Coefficient respectively. From these results, it is observed that the proposed approach achieved a reduced computational time, less false alarm rate, and high correlation coefficient.

5. Conclusion

This paper proposed an efficient IDS aimed at reducing the computational time and to enhance the accuracy. The proposed system is built based on a new Multi Linear Dimensionality Reduction through which the computational time decreases in an efficient manner. The proposed ML-DR removes the noise attributes and retains the optimal attribute set. Multi-Class SVM constructs classification model based on the training data obtained from ML-DR. Optimization of SVM parameters for the RBF kernel is carried out through PSO. NSL-KDD dataset was used for performance evaluation. The experimental results indicate that the classification accuracy of proposed approach outperforms the other classification techniques.

The proposed method will significantly reduce both the memory size and the CPU time required for intrusion detection by grouping all the features used for the detection. This shows that the proposed method is very reliable for intrusion detection. Results indicate that the proposed Multi Linear Dimensionality Reduction detection method outperforms other methods since it can provide better and more robust representation of the data. This is due to the fact that it can accurately detect a broader range of attacks using smaller number of features. In this paper, we compare the performances of different methods used for Intrusion detection on the basis of DR and FAR values. The results obtained as, the value of rates obtained from Genetic algorithm is quite lesser than Multi Class SVM. And increased efficiency of Multi Linear Dimensionality reduction with Multi-Class SVM is increased up to (98% approx.) much better than the two approaches which is up to (95% approx.).

In order to achieve still more accurate results, the proposed approach can be accomplished in two phases by means of two stage classification. The first stage classification performs an unsupervised classification though clustering the entire dataset and followed by Multiclass Support vector machine classifier at the second phase.

References

[1] S. Axelsson, “Research in intrusion-detection systems: a survey”, *Technical report*, 1998.

[2] M. Tavallaee, A.A. Ghorbani, and W. Lu, “Network Intrusion Detection and Prevention: Concepts and Techniques”, *Advances in Information Security, Springer Science (eBook)*, Vol.47, No.1, pp.1-216, 2009.

[3] P. García-Teodoro, “Anomaly-based network intrusion detection: Techniques, systems and challenges”, *Computers & Security*, Vol. 28, No.2, pp.18-28, 2009.

[4] A. Sung and S. Mukkamala, “Identifying important features for intrusion detection using support vector machines and neural networks”, In: *Proc. of the International Symposium on Applications and the Internet (SAINT 2003)*, pp. 209–217, 2003.

[5] Z. Safaa, and K. Fakhri, “Lightweight IDS based on feature selection and IDS classification scheme”, In: *Proc. of International Conf. on Computational Science and Engineering*, pp. 365–370, 2009.

[6] F. Amiri, M.M.R. Yousefi, C. Lucas, A. Shakery, and N. Yazdani, “Mutual information based feature selection for intrusion detection systems”, *Journal of Network Computer Applications*, Vol. 34, No. 4, pp.1184–1199, 2011.

[7] S. Balakrishnan, K. Venkatalakshmi, and A. Kannan, “Intrusion detection system using feature selection and classification technique”, *International Journal of Computer Science and Application (IJCSA)*, Vol. 3, No. 4, pp.145–151, 2014.

[8] H. Saxena, and V. Richariya, “Intrusion detection inKDD99 dataset using SVM-PSO and feature reduction with information gain”, *International Journal of Computer Applications (IJCA)*, Vol. 98, No. 6, pp. 25–29, 2014.

[9] S.V. Farrahi, and M. Ahmadzadeh, “KCMC: A hybrid learning approach for network intrusion

detection using K-means clustering and multiple classifiers", *International Journal of Computer Applications,* Vol. 124, No. 9, pp. 18–23, 2015.

[10] G. Liu, "An intrusion detection model based on the PCA and neural networks", *Neurocomputing,* Vol. 70, Issue 7-9, pp. 1561–1568, 2007.

[11] F. Kuang, "A novel hybrid KPCA and SVM with GA model for intrusion detection", *Applied Soft Computing,* Vol. 18, pp. 178–184, 2014.

[12] S. Thaseen and C. A. Kumar, "Intrusion Detection Model using fusion of PCA and optimized SVM", In: *Proc. of International Conf. on Computing and Informatics (IC3I),* Mysore, India, pp. 879–884, 2014.

[13] S.T. Ikram and C. A. Kumar, "Intrusion Detection Model using fusion of chi-square feature selection and multi class SVM", *Journal of King Saud University – Computer and Information Sciences,* Vol.29, pp.462-472, 2017.

[14] S.T. Ikram and C. A. Kumar, "Intrusion Detection model using chi-square feature selection and modified naïve Bayesian classifier", In: *Proc. of third International Symposium on Big Data and Cloud Computing challenges,* pp.81-91, 2016.

[15] B. Kasliwal, S. Bhatia, S. Saini, I.S. Thaseen, and C.A. Kumar, "A hybrid anomaly detection model using G-LDA", In: *Proc. of IEEE International Advance Computing Conference,* pp. 288–293, 2014.

[16] S. Ganapathy, *"*A Novel Weighted Fuzzy C-means Clustering Based on Immune Genetic Algorithm for Intrusion Detection", *Procedia Engineering,* Vol. 38, pp.1750-1757, 2012.

[17] A.J. Malik, W. Shahzad, and F.A. Khan, "Binary PSO and random forests algorithm for probe attacks detection in a network", In: *Proc. of IEEE Congress on Evolutionary Computation (CEC),* pp. 662–668, 2011.

[18] D. K. Srivastava and L. Bhambhu, "Data classification using support vector machine", *Journal of Theoretical and Applied Information Technology,* Vol. 12, No. 1, pp. 1–7, 2010.

[19] C.W. Hsu, C.C. Chang, and C.J. Lin, "A Practical Guide to Support Vector Classification", *Department of Computer Science National Taiwan University, Taipei 106, Taiwan,* 2016.

[20] R. Chen, D.Y. Sun, D.T. Qin, and F.B. Hu, "A novel engine identification model based on support vector machine and analysis of precision-influencing factors", *Journal of Central South University (Science and Technology),* Vol. 41, No. 4, pp.1391–1397, 2010.

[21] M. Tavallaee, E. Bagheri, W. Lu, and A.A. Ghorbani, "A detailed analysis of the KDD CUP 99data set", In: *Proc. of the Second IEEE Symposium on Computational Intelligence for Security and Defense Applications,* 2009.

[22] M. Panda, A. Abraham, and M.R. Patra, "Discriminative multinomial naive Bayes for network intrusion detection", In: *Proc. of Sixth International Conf. on Information Assurance and Security (IAS),* pp. 5–10, 2010.

[23] J. McHugh, "Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory", *ACM Transactions on Information and System Security,* Vol. 3, No.4, pp.262–294, 2000.

[24] T.R. Reddy, B.V. Vardhan, and P.V.P. Reddy, "A Survey on Authorship Profiling Techniques", *International Journal of Applied Engineering Research,* Vol.11, No.5, pp.3092-3102, 2016.

[25] Q. Yang, H. Fu, and T. Zhu, "An Optimization Method for Parameters of SVM in Network Intrusion Detection System", In: *Proc. of International Conf. on Distributed Computing in Sensor Systems,* Washington, DC, USA, pp.136-142, 2016.