International Journal of Engineering & Technology, 7 (2.7) (2018) 277-283

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Intrusion detection mechanism with machine learning process A case study with FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian methods

K V S S R Murthy [1] *, K V V Satyanarayana [2]

[1] Research Scholar, Department of CSE, K L Educational Foundation, Vaddeswaram, Guntur, AP-522502, India
[2] Professor, Department of CSE, K L Educational Foundation, Vaddeswaram, Guntur, AP-522502, India
**Corresponding author E-mail: kvssrmurthy75@gmail.com*

Abstract

Today, there is a far reaching of Internet benefits everywhere throughout the world, numerous sorts and vast number of security dangers are expanding. Since it isn't in fact possible to assemble a framework without any vulnerability, Intrusion Detection System (IDS), which can successfully distinguish Intrusion, gets to have pulled in consideration. Intrusion detection can be characterized as the way toward distinguishing irregular, unauthorized or unapproved action that objective is to target a system and its assets. IDS plays a very important role for analyzing the network passage, also it assumes a key part to analyze the system activity log and each log is portrayed by huge arrangement of highlights and it requires tremendous computational preparing force and time for the characterization procedure. This work proposes filter based feature selection methods to predict intrusion with Feature based Mutual Information Feature Selection Support Vector Machine (FMIFSSVM), Feature based Liner Correlation Feature Selection Support Vector Machine (FLCFSSVM), misuses SVM, anomaly SVM and Bayesian methods. The performances of these methods are considered by using the intrusion detection calculation data set called Knowledge Discovery in Databases (KDD) cup 99. Detection Rate (DR), False Alarm Rate (FAR) and Percentage of Successful Prediction (PSP) are the major performance measures studied in this work.

Keywords: *IDS; FMIFSSVM; FLCFSSVM; Misuse SVM; Percentage of Successful Prediction (PSP).*

1. Introduction

The art of determining the activities over the network which is not as per the predefined security policies is major role of Intrusion detection system [1]. An ID is somewhat high-quality ability of the methods for intrusion detection that are developed in the earlier decades. Firstly Intrusion detection systems will help to the nodes in the network to make them ready to react to network attacks by attackers.

Intrusion detection system will generally responsible for the following activities:

- Observing the behavior of both client and node activities.
- Observing the system configurations and effects.
- Determining the system and integrity of the files.
- capability to make out patterns distinctive attacks
- investigation of nonstandard action patterns
- Finding the actions against the allowed user behaviors.

The motivation in the wake of IDS is to assist PC on the highly accomplished method to manage activities, and all this is done by IDS which gather entire required data from one or more special sources and systems , later these are compared with some other predefined patterns of violations to decide whether these are attacks or security violations.

Malicious behavior is defined as any human being action which try to make use of or get accessed to a node[2] without taking any permission of other human being who have legitimate access to the node.

The word attack can be defined as sequence of actions done by a human being or node to make damage to the system or entire network [3]. Based on the security violations of some feature values attack type is defined. It was also observed that some intruders determine the relationship between attack type and a particular node in the network and from that also they take the advantage. So we should concentrate on unseen relationships between attack types and nodes in the network, this will help us in identifying some type of attacks. Some other attacks generally begins with an effort to access the gain over some node in corporate network through some resource and place supplementary or advanced actions on the network, that's why we have to define the attack prediction process[4] as a series of primary actions that are applied to determine the attack strategy. The attackers make use of distributed and coordinated techniques, as a result it becomes more difficult to find that a particular action is attack or not. The information is collected from different un matched events and from different corporate networks and combine this information together to reveal the attack plan. Therefore it is necessary to design a more advanced prediction system to deal with different attack strategies, so that correct remedies and measures are taken in advance to decrease the effect of attack and to stay away from some likely attacks. Now this work proposes an anti intrusion system should concentrate on uncertainty that can control the information. Generally there are two sources of uncertainty [5], one is stochastic uncertainty and other is epistemic uncertainty. The stochastic uncertainty also called variability it is the result of unconfident nature or random information that is appropri-

ate to a likely inconsistency. And the other one epistemic uncertainty is the result of partial or incorrect or not having complete knowledge and availability of the data.
Denial of service attacks will always try to smash the nodes in the network by observing the traffic over the network [1]. The major aim of network IDS to detect this type of illegal activities. The major problem in this domain is to build a representation for network intrusion detection and make exercise it for further identifying and made an investigation of intrusions. To solve this problem a good number of attempts are proposed. Some of the current methods could solve up to some extent, but there is a lot of scope to further improvement. And it is also observed that no single method provides solutions to all the identified problems.
The major challenge field of network intrusion detection is to shape the area of intrusion and to reason on the model. Knowledge engineering approach and data mining approach [6] are the basic flavors of designing a domain model in IDS. All the people who are experts in that domain in along with one or more knowledge engineers and recognize the relations connecting domain variables by general manual approach. That's why this method of knowledge engineering approach is suitable when we have less no of variables and they can control manually. So, it is not sensible when there is more number of variables, and even several times, it is totally unreasonable as some problems are most difficult and they have more no of variables. And the other method called data mining approach ,in this model the area model is obtained robotically from an algorithm, for this algorithm the input is network intrusion data and output is the suitable model. In this way this method of finalizing the model does not need human effort. Nevertheless, the victory of data mining approach is always depends up on, how we are accessing the large dataset. Data mining approach is truly appropriate to detection of some well defined intrusions [7]. The KDD Cup 1999 data set is the familiar to test the intrusion over the network, as it consists of 41 different variables and also having marvelous 4 million records [8]. Underlying logic is all about identifying the origin and analyses the outcome of the intrusion. Here in this paper, a general knowledge-driven method is taken to determine the intrusions. The authors also said that this knowledge driven method is suitable for medical and mechanical fields where as this method is less used while detecting and analyzing the intrusions.
To test whether the IDS works correctly or not a standard database is needed, for this KDDCUP99 dataset is downloaded. The KDDCUP99 dataset has a good range of network traffic records. In KDDCUP99 there are two datasets one complete dataset 743M uncompressed and the other is 10% subset approximately 75M uncompressed, the second one is used as tentative dataset.
To link the gap, FMIFSSVM, FLCFSSVM, misuses, anomaly SVM and Bayesian methods are planned. It deals with the modeling and reasoning problems. The combination of a good machine learning algorithm and Bayesian network serves as better modeling approach in the area of IDS. This combination gives efficient reasoning capability. It was also observed that the spending and human effort in construction of Bayesian causal model is reduced by the usage of machine learning techniques.

2. Literature survey

Warrender et.al [9] proposed some IDS techniques with system call trace data by sliding window concepts. In this authors form testing database for instance testing. They tested this proposed method with related techniques. They also categorize the testing instances based on the database of normal sequence.
Wenke Lee et.al [10] presented an offline IDS for networking and host systems, in which they identifying the anamoly and misuse IDS based on association rules.
Agarwal R.et.al[11] projected a two stage structure to learn a rule based model that studied the classifier models. For testing this proposed method authors used the KDDCUP99 dataset.
Yinhui Li et al [12] observed that the dimension of data features will also give success in the detection of intrusions.
He found that Support vector machine SVM is proved its strength and competence in the network classification and so SVM is also useful in determining intrusion detections. According to Yinhui Li, SVM, clustering method and ant colony algorithms together will determine the network traffic is allowed or disallowed. While testing this they take help of KDDCUP99 dataset.
D. Barbara et. al. [13] considers a data set which does not contains any attacks and then used an algorithm to grouping of attacks with known and unknown behavior and false alarms.
T. Abraham [14] develops an Intrusion detection system for misuses and anomaly detection using characteristic of meta rules and assosiciation rules, It also involves data mining to create report about network traffic and use this report for deviation analysis.
Neural network classifier based on statistical methods for anomaly detection was projected by Z. Zhang et. al. [15] to recognize UDP flood attacks. they also stated that back propagation neural network (BPN) is also very useful in developing competitive IDS.
A. Chauhan et. al. [16] explains the use of data mining techniques in IDS and also brief the pros and cons of different data mining techniques.
Xin Xu et. al. [17] explains how the combination of Multiclass SVM and machine learning techniques are useful in developing IDS.
Yang Li et. al. [18] projects the inabilities of KDDCUP99 dataset and proposed a most lively machine learning named TCM-KNN.
M.Panda et.al [19] enhances the ID3, J48 and naïve Bayes classifying techniques. For testing this proposed method authors used the KDDCUP99 dataset.
Mohammed M. et. al. [20] mentions the ruthlessness of attacks over the network with the support of the comprehensive analysis classification techniques. This paper also describes the features of Z- R classifier along with decision table classifier with the support of KDDCUP99 dataset in MIT Lincoln laboratory.
S. Sathyabama et. al. [21] also decides the user's activities or according to security policies or not and call the un secured activities as outliers, for this they proposed clustering techniques.
Ayei et. al. [22] discusses both misuses and anomaly detection methods while discovering a new cross technique by combining the ideas of J48, Boyer Moore and K-NN algorithms so that it detects the intrusions better, this algorithm is HYBRITQ-4.

3. Methodology

In data mining we should find all the constructive data and reasonable data from bulky data after applying proper algorithm. It uses of a number of algorithms or methods in a way to unite with statistics, artificial intelligence, machine learning, database science, and information retrieval [23].
In outline data mining is:

- a group of techniques that will find patterns in data
- it is very useful to the users, which improves analysis of data and computing authority
- it can be a collection of techniques that are used to find relationships that are not discovered earlier.
- It does not dependent on an present database

Data mining consists of following key essentials:

- It is responsible for Extracting data from the data warehouse, transform it to required form and again reload the transactional data onto the data warehouse system.
- It is responsible for Storing and managing data in both two dimensional as well as multidimensional database system.
- Mainly it gives access to business experts and data modernization experts.
- It Analyze information by application training software.
- It Present information in a supportive arrangement, like tables, bar charts or pie charts.

Role of Data Mining in IDS: to get better IDS [24], the centre of attention is anomaly detection; here after finding the limits of the network traffic model is identified. The well known dynamic models are Supervise learning and unsupervised learning.

- Supervised learning gain knowledge of models from history of intrusions
- Unsupervised learning: will identify the new mistrustful behaviors

Data mining along with machine learning techniques will assist in differentiate suspicious actions from most natural daily network passage.

Data mining deals with the following familiar tasks.

Anomaly detection: it means the detection of strange information or data records which are affecting or statistical errors require additional examination.

Association rule learning: this is also called dependency modeling during this searching for associations are done among variables. As an example a store might collect data on client purchasing behavior. And based on this information it apply association rule learning, as a result the store will get information such as the items that are commonly bought jointly and they can re use this data for advertising purpose. This is now and then called as market basket analysis.

Clustering: in the clustering process we find the similar groups and structures present in the relevant information.

Classification: classification is the work of generalize well known configuration to concern to innovative data. Classification works in a similar way as an email series has ability to categorize whether the particular email as authentic or spam.

Regression: the aim of regression to find a particular function to model the available data by means of the slightest error.

Summarization: it includes visualization and report generation.

3.1. Feature selection

During Feature selection process we identify and remove in appropriated and duplicate features to determine most favorable subset of features that construct enhanced characterization of patterns relating to dissimilar classes.

Filter method and Wrapper method [25] are most general form of feature selection. Filter method algorithms will estimate the relation between a set of features based on autonomous measures like distance measures, consistency measures and information measures. Whereas wrapper method algorithms will estimate the value of features based on some learning algorithms. When we consider the big enough datasets, the filter methods are computationally feasible and wrapper methods are computationally expensive. This study focus on filter methods of IDS.

3.2. Mutual information

Mutual Information can handle both linearly dependent and non linearly dependent data and is competently measures the variable dependence estimation [25]. The study of probability theory concludes that the mutual information of two arbitrary variables is a measure of mutual dependence among those variables. It was already known from earlier studies that mutual information is always related to entropy of a arbitrary variable [25]. Most generally Mutual information will find the likeness between joint distribution p(X, Y) and multiplication of factored marginal distribution of p(X)p(Y).

More generally, while considering the discrete random variables X and Y then their MI can be given in the following equation (1)

$$I(X;Y) = \sum_{y\in Y}\sum_{x\in X} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (1)$$

In the above equation (1) p(x,y) denoted the joint probability function of X and Y, and p(X)p(Y) will denoted the marginal probability distribution functions of X and Y correspondingly.

When the continuous random variables are considered, the summation is changed to double integral; this is given in the equation (2).

$$I(X;Y) = \int_Y \int_X p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right)dxdy \quad (2)$$

In the above equation (2) p(x,y) will denote the joint probability density function of X and Y, and p(x) and p(y) will denote the marginal probability density functions of X and Y correspondingly. Again as we use the log base 2, the units of mutual information will be bits.

3.3. Support vector machines

SVM are considered as supervised learning models that can map actual input characteristic vector to better dimensional characteristic space by using non linear mapping [27] SVM know how to provide real-time discovery ability while dealing with huge dimensionality of data. SVM classifies data with the help of hyper plane in the characteristic space. This procedure will lead to the problem of quadratic programming.

For example consider N educational information points

{(x1, y1), (x2,y2), (x3, y3), ..., (xN , yN)},

Where x¡ € Rd and y¡ €{+1,−1}. Let a hyper-plane is defined by (w, b), where w is a weight vector and b is a bias, then the classification of a new object x is done with N given away in equation 3.

$$f(x) = sign(w.x + b) = sign(\sum_{\alpha_i y_i}(x_i.x) + b) \quad (3)$$

Here the training vectors xi appears merely in the shape of a dot product.

We consider the Lagrangian multiplier say $\alpha_¡$ corresponding to every preparation data point. The Lagrangian multiplier values like α_i give the significance of every preparation data point.

It was also found that the only the points which lie nearer to hyper plane are having $\alpha_i > 0$ means the maximal margin hyper-plane is found, and such points become the support vectors, and remaining points are having $\alpha_i = 0$. Which conclude that all the points which are nearer to hyper plane will always lead to hypothesis/ classifier?

3.4. Bayesian network classifier

The Bayesian classifier [28] is the most successful classifier. When we compare Bayesian classifier with others, its predictive performance is always superior.

Consider the exercised training data, the conditional probability of the attribute A_i is specified in the class label C. by the application of Bayes rule the Classification to this class can be done to calculate the probability of class labeled C known the some familiar individual instance of A1; : : : ; An, now the classifier will predict the class with the maximum subsequent Probability.

This calculation is always possible by construction a tough independent statement that all the different attributes Ai are conditionally independent specified the value of the class labeled C. the word independent here means probabilistic independence, or in other words, A is independent of B C whenever $Pr(A_jB;C) = Pr(A_jC)$ for all likely values of A;B and C, whenever $Pr(C) > 0$ [10].

Bayesian networks are the symbol of a innovative move towards finding or avoidance of attacks in valuable networks; we can use Bayesian networks while identifying attacks through IDS to solve the most commonly raised troubles in earlier methods. Bayesian networks have major advantages which are not likely to realize by means of other methods. Network guidance is not achievable in a genuine atmosphere; so, systems for intrusion detection or intrusion prevention are not open. The earlier knowledge on attacks of entire network or individual nodes is all integrated in one system called IDS. IDS can also estimate the probability of attacks when events can happen.

IDS System with Bayesian network will provide a special advantage over other system with no Bayesian network, when we calculates the power of newly shaped events with the earlier events. For that reason, entire data and regulations can be used in new systems; these can be structured into IDS with the support of Bayesian networks. Bayesian networks give a complete compatibility of equivalent software components. Without considering to policy

used for implementation; and so accelerating the growth and function of individual IDS or distributed IDS.

3.5. About NIDS architecture

We recommend a Network intrusion detection system NIDS, shown in Fig. 1, having four major parts namely packet preprocessing, intrusion detection (both and anomaly detection and signature based), storage which comprises central log ,knowledge base, behavior base then the final one is alert system.

Snort and FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers are the sub parts of intrusion detection. The purpose of snort is to distinguish already recognized attacks after correlating them with captured packets with some novel strategy in the knowledge base. FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers will first find the class label of already preprocessed packets. The module named Storage comprises of the databases namely central log, knowledge base and behavior base. The purpose of Knowledge base is to stock up all the recognized attack signatures, while the purpose of behavior base to stock up the network actions having cruel packets separates them with normal packets and the purpose of Central log is to register cruel incident that is measured by FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers.

The modules of NIDS are deployed on additional servers to fill in their knowledge base and also in the behavior base, whenever the central log is updated. While the purpose of Alert system is to create an alert, whenever some conflict is identified by FMIFSSVM, FLCFSSVM misuses SVM, anomaly SVM and Bayesian classifiers.

The work flow of NIDS is depicted in Fig. 2, network packets are gathered from outside or inside of the networks. Snort then compares the captured packets with the system policies that are already located within knowledge base. If any relationship is identified immediately an alert be signaled, and then place the information in the central log.

The packets which do not have any coincidence are called Normal packets; these are preprocessed and forwarded to FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers, these classifiers will predicts the group label of those packets.

FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian are skilled by means of behavior base. If FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers identifies some intrusions, those will be alerted and immediately they are positioned in the central log or else, those individual packets are treated as normal actions.

In the NIDS, the combination of anomaly and signature based detection is used which improves the correctness of intrusion detection ratio. Here first the signature based detection technique is applied and then anomaly detection is applied, as result the computational cost becomes less. FMIFSSVM, FLCFSSVM, misuses SVM, anomaly SVM and Bayesian classifiers are targeted to identify only unidentified and mysterious attacks, since the earlierly recognized attacks are previously identified in the unit Snort. NIDS can be located on any of the servers to modernize their knowledge base by receiving alerts placed in central log, consequently, any mysterious attack which was earlier identified at any one of the server be able to be effortlessly identified by Snort at additional servers. In this way the computational costs meant for identifying mysterious attacks at new servers is decreased drastically.

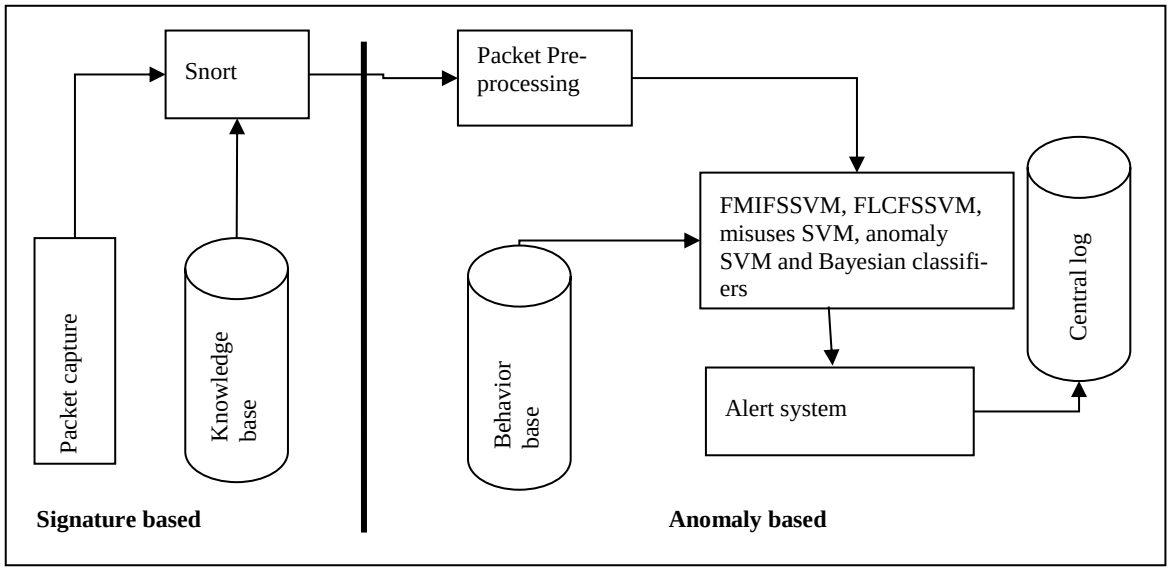

Fig. 1: Proposed NIDS Module.

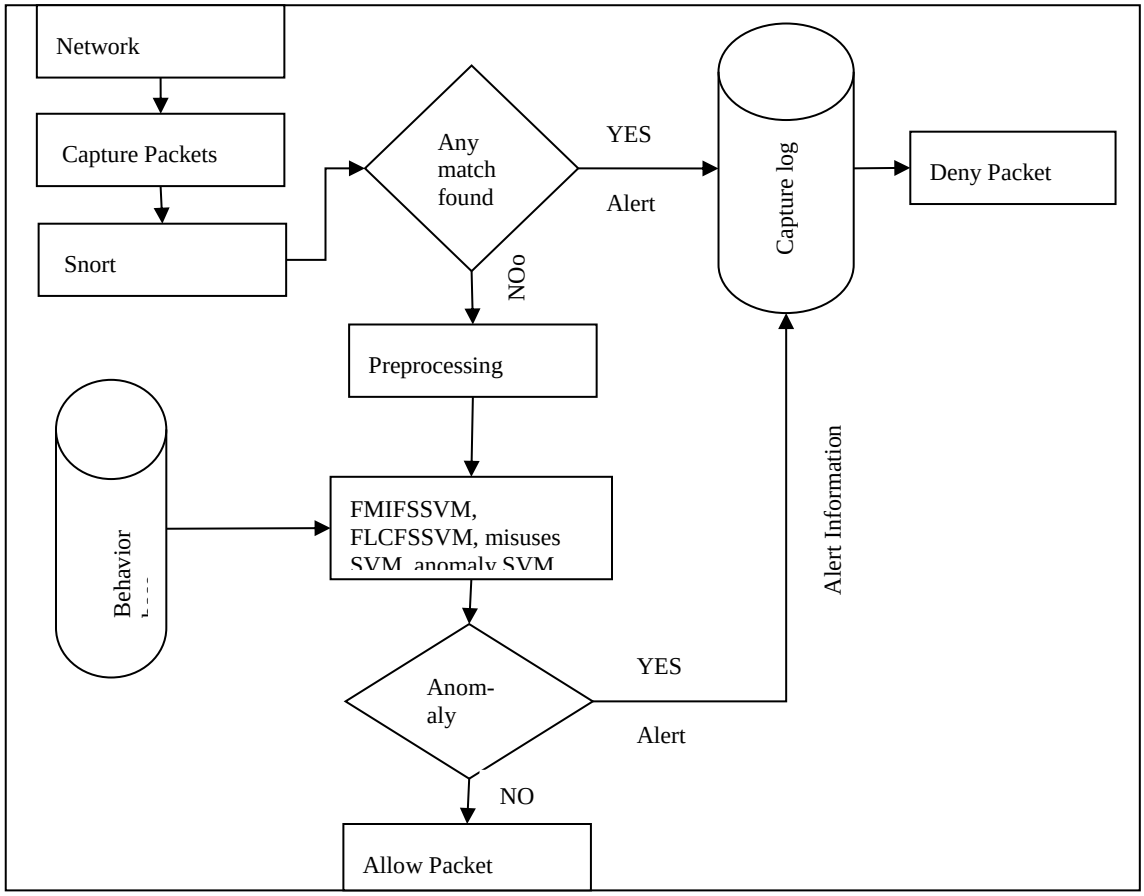

Fig. 2: Workflow of NIDS Module.

4. Performance comparison measures

To make analysis on KDDCUP99 training dataset we can use special types of data mining techniques, these tools have their own advantages and disadvantages. Here, five methods namely FMIFSSVM, FLCFSSVM, misuse SVM, anomaly SVM and Bayesian are considered and these methods are tested by some sample data from KDD 99 dataset. For testing FMIFSSVM, FLCFSSVM methods, we take 192 samples from KDD 99 dataset. For testing misuse SVM, anomaly SVM methods, we take 195 samples from KDD 99 dataset. For testing Bayesian methods, we take 11850 samples from KDD 99 dataset. These statistics are shown in Table 1.

To grade these consequences, the standard metrics Detection Rate (DR) and False Alarm Rate(FAR) have already been used; as in Equations 4, 5, 6, 7, 8 and 9 respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad (4)$$

$$Detection\ Rate = \frac{Number\ of\ samples\ classified\ correctly}{Number\ of\ samples\ used\ for\ traininig}\ or\ \frac{TP}{TP+FN} \quad (5)$$

$$False\ Alarm\ rate = \frac{False\ positives}{total\ number\ of\ normal\ connections} \quad (6)$$

$$False\ Positive\ Rate = \frac{FP}{FP+TN} - - - \quad (7)$$

$$F - measure = \frac{2(precision*recall)}{precision+recall} \quad (8)$$

In the above equations

TP = True Positive it denotes the quantity of real records recognized as attacks.

TN = True Negative it denotes quantity of real usual records recognized as usual ones.

FP = False Positive it denotes quantity of real usual records identified as attacks.

FN= False Negative it denotes quantity of real usual records recognized as normal records.

Table 1: Detection Rate (Dr) And False Alarm Rate (Far)

METHOD NAME	Data Set Size	Detection Rate (DR)	False Alarm Rate(FAR)
FMIFSSVM	192	84.89	20.94
FLCFSSVM	192	84.89	25.87
Misuse SVM	195	90.90	28.79
Anomaly SVM	195	78.79	27.27
Bayesian	11850	60.18	12.54

From the Table I, it is observed the Detection Rate (DR) 84.89 for FMIFSSVM and FLCFSSVM of 192 samples which is for correct predictions --163/total--192=84.89. It is also observed False Alarm Rate (FAR) 20.94 for both FMIFSSVM and FLCFSSVM of 192 samples.

From the Table I, it is also observed the Detection Rate (DR) 90.90 for misuse SVM of 195 samples which is for correct predictions --177/total--195=90.90. It is also observed False Alarm Rate (FAR) 28.79 for misuse SVM of 195 samples.

From the Table 1, it is also observed the Detection Rate (DR) 78.79 for anomaly SVM of 195 samples which is for Correct predictions --153/total--195=78.79. It is also observed False Alarm Rate (FAR) 27.27 for anomaly SVM of 195 samples.

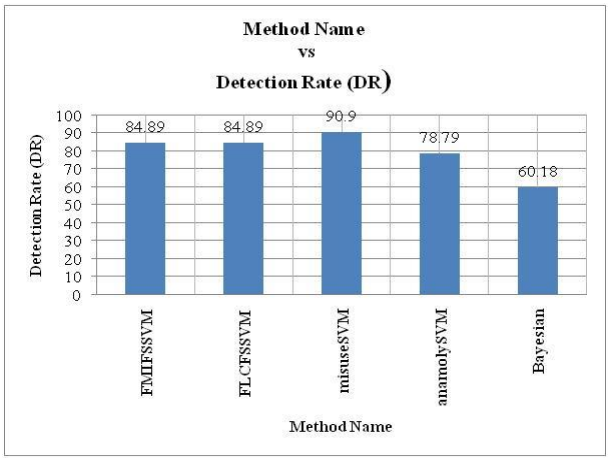

Fig. 3: Detection Rate (DR) for FMIFSSVM, FLCFSSVM, Misuse SVM, Anomaly SVM and Bayesian Methods.

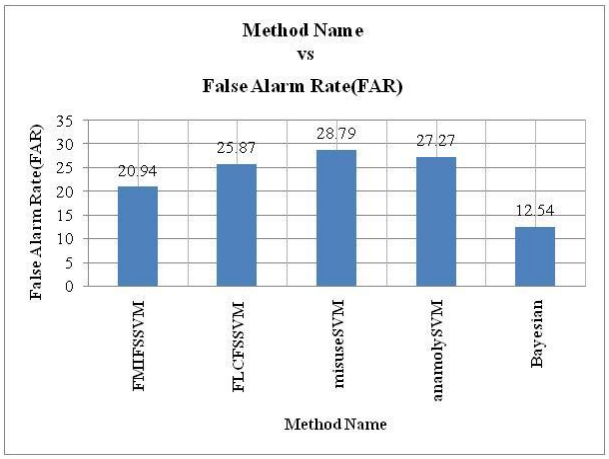

Fig. 4: False Alarm Rate (FAR) for FMIFSSVM, FLCFSSVM, Misuse SVM, Anomaly SVM and Bayesian Methods.

From the TABLE I, it is also observed the Detection Rate (DR) 60.18for Bayesian of 11850samples which is for Correct predictions --7132/total--11850=60.18. It is also observed False Alarm Rate (FAR) 12.54 for Bayesian of 11850 samples.

Confusion Matrix is an additional quantity of performance of IDS it consists of elements like actual connection label and predicted connection label as shown in TABLE II.

The graphical representation of Detection Rate (DR) for FMIFSSVM, FLCFSSVM, misuse SVM, anomaly SVM and Bayesian methods is shown in Figure 3.

The graphical representation of False Alarm Rate (FAR) for FMIFSSVM, FLCFSSVM, misuse SVM, anomaly SVM and Bayesian methods is shown in Figure 4.

Table 2: Standard Metrics for Evaluation of Intrusions (Attacks) in KDD 99 Dataset

Confusion Matrix (Standard Metrics)		Predicted Connection Label	
		Normal	Intrusions(Attacks)
Actual Connection label	Normal	True Negative(TN)	False Alarm(FP)
	Intrusions(Attacks)	False Negative (FN)	Correctly detected Attacks(TP)

Table 3: Percentage of Successful Prediction (PSP)

METHOD NAME	Data Set Size	Percentage of Successful Prediction (PSP)
FMIFSSVM	192	84.89
FLCFSSVM	192	84.89
misuses	195	90.90
anamolySVM	192	78.79
Bayesian	192	60.18

The equation 10 precise that the Percentage of Successful prediction (PSP) will lead to accuracy

$$PSP = \frac{Number\ of\ incidents\ that\ have\ been\ successfully\ classified}{Total\ number\ of\ incidents} \times 100\% \quad (10)$$

When the PSP is high the classification for IDS is better and the low PSP gives low classification for IDS is low. Here PSP is used to rank different methods like FMIFSSVM, FLCFSSVM, Misuses SVM, anomaly SVM and Bayesian these statistics are shown in TABLE III.

From the Table III, it is observed the Percentage of Successful Prediction (PSP) 84.89 for both FMIFSSVM & FLCFSSVM of 192 samples. It is also observed the Percentage of Successful Prediction (PSP) 84.89 for misuse SVM of 195 samples. It is also observed the Percentage of Successful Prediction (PSP) 78.79 for anomaly SVM of 195 samples, similarly the Percentage of Successful Prediction (PSP) 60.18 for Bayesian of 11850 samples. It is also observed the Percentage of Successful Prediction (PSP) 60.18 for Bayesian of 11850 samples. The graphical representation of Percentage of Successful Prediction (PSP) for FMIFSSVM, FLCFSSVM, misuse SVM, anomaly SVM and Bayesian methods is shown in Figure 5.

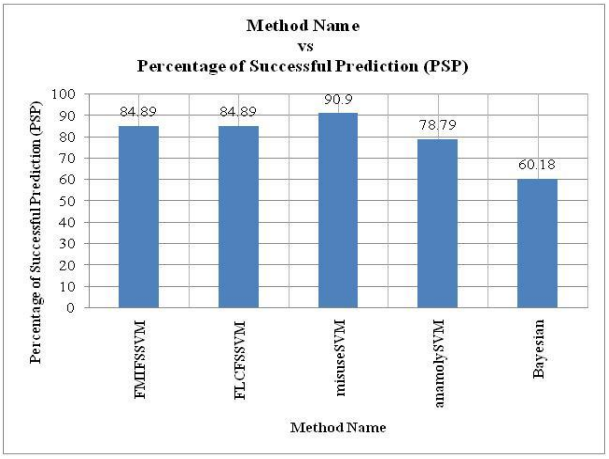

Fig. 5: Percentage of Successful Prediction (PSP) for FMIFSSVM, FLCFSSVM, Misuse SVM, Anomaly SVM and Bayesian Methods

5. Conclusions

The aim of this effort was to present the correctness of the a variety of classes of attacks using FMIFSSVM, FLCFSSVM, misuse SVM, anomaly SVM and Bayesian methods, we have achieved our target by implementing these methods as an engine to categorize the data as a result. Comparisons has been made among these classifiers and the results are shown in Table 1 & 2 and also in Figures 1, 2 & 3. These results achieved for some of the samples from KDD 99 dataset. Although these methods have shown hopeful in PSP, it might be enhanced by optimizing this IDS strategy to large datasets.

References

[1]. Amiri M, Feizi-Derakhshi MR, Mohammad-Khanli L, "IDS fitted Q improvement using fuzzy approach for resource provisioning in cloud", Journal of Intelligent & Fuzzy Systems, Vol.32, issue.1, pp.229-240, 2017.

[2]. Pacheco J, Zhu X, Badr Y, Hariri S., "Enabling Risk Management for Smart Infrastructures with an Anomaly Behavior Analysis Intrusion Detection System", International Workshops on Foundations and Applications of Self* Systems (FAS* W), pp. 324-328, 2017.

[3]. Puri A, Sharma N., "A novel technique for intrusion detection system for network security using hybrid SVM-cart", International Journal of Engineering Development and Research, Vol.2, issue,17, pp. 155-161, 2017.

[4]. Stefanova Z, Ramachandran K., "Network attribute selection, classification and accuracy (NASCA) procedure for intrusion detection systems", International Symposium on Technologies for Homeland Security (HST), pp. 1-7, 2017.

[5]. Han J, Qian C, Wang X, Ma D, Zhao J, Xi W, Jiang Z, Wang Z., "Twins: Device-free object tracking using passive tags", IEEE/ACM Transactions on Networking, vol.24, issue.3, pp.1605-1617, 2016.

[6]. Sharma N, Gaur B., "An approach for efficient intrusion detection for KDD dataset: a survey." International Journal of Advanced Technology and Engineering Exploration, vol.3 (18), pp: 72. 2016.

[7]. Kumar GR, Mangathayaru N, Narsimha G.," Intrusion Detection-A Text Mining Based Approach." International Journal of Computer Science and Information Security. Feb 1, vol.1, pp.76, 2016.

[8]. Ruan Z, Miao Y, Pan L, Patterson N, Zhang J., "Visualization of big data security—A case study on the KDD99 cup data set:. Digital Communications and Networks. Aug 12.,2017

[9]. Warrender C., Forrest S. and Pearl M.,"Detecting Intrusions Using System Calls: Alternative Data Models", in IEEE symposium on security and privacy, pp:133-145, 1999.

[10]. Wenke L. and S. J.Stolfo, "A Framework for Constructing Features and Models for Intrusion Detection Systems", ACM transactions on Information and system security (TISSEC), vol.3, Issue 4, Nov 2000.

[11]. Agarwal R., Joshi M.V., "PNrule: A New Framework for Learning Classifier Models in Data Mining", Tech. Report, Dept. of Computer Science, University of Minnesota, 2000.

[12]. Li Y, Xia J, Zhang S, Yan J, Ai X, Dai K.," An efficient intrusion detection system based on support vector machines and gradually feature removal method". Expert Systems with Applications.vol.1, pp. 424-30, 2012.

[13]. Daniel B., J.Couto, S.Jajodia, and N.Wu, "ADAM: A Test Bed for Exploring the Use of Data Mining in Intrusion Detection", SIGMOD, vol30, issue no.4, pp: 15-24, 2001.

[14]. Abraham T., "IDDM: Intrusion Detection Using Data Mining Techniques", Technical report DSTO electronics and surveillance research laboratory, Salisbury, Australia, May 2001.

[15]. Zheng Z., J. Li, C.N. Manikapoulos, J.Jorgenson, J.ucles, "HIDE: A Hierarchical Network Intrusion Detection System Using Statistical Pre-Processing and Neural Network Classification", IEEE workshop proceedings on Information assurance and security, pp: 85-90, 2001.

[16]. Yeung D. Y. and Chow C., "Prazen-window Network Intrusion Detectors", in: 16th International Conference on Pattern Recognition, Quebec, Canada, pp: 11–15, 2002.

[17]. Xu X., "Adaptive Intrusion Detection Based on Machine Learning: Feature Extraction, Classifier Construction and equential Pattern Prediction", International Journal of Web Services Practices vol 2, issue 1, pp: 49–58, 2006.

[18]. Li Y., Guo L., "An Active Learning Based TCM-KNN Algorithm for Supervised Network Intrusion Detection", In: 26th Computers and Security, pp: 459–467, 2007.

[19]. Mrutyunjaya P. and M. Ranjan Patra," Evaluating Machine Learning Algorithms for Detecting Network Intrusions", International Journal of Recent Trends in Engineering, vol. 1, no.1, 2009.

[20]. Mohammed M Mazid, M. Shawkat Ali, Kevin S. Tickle,"A Comparison Between Rule Based and Association Rule Mining Algorithms ", Third International Conference on Network and System Security, 2009.

[21]. Sathyabama S., Irfan Ahmed M., Saravanan A,"Network Intrusion Detection Using Clustering: A Data Mining Approach", International Journal of Computer Application (0975-8887), vol. 30, no. 4, Sep. 2011.

[22]. Ibor AE, Epiphaniou G." A Hybrid Mitigation Technique for Malicious Network Traffic based on Active Response", International Journal of Security and Its Applications. Vol.9,issue 4,pp: 63

[23]. da Cunha JA, Moura E, Analide C. "Data Mining in Academic Databases to Detect Behaviors of Students Related to School Dropout and Disapproval", InWorldCIST (2) pp. 189-198, 2016.

[24]. Buczak AL, Guven E. "A survey of data mining and machine learning methods for cyber security intrusion detection". IEEE Communications Surveys & Tutorials. vol 1;issue 2:pp.1153-76, 2016

[25]. Novaković J. "Toward optimal feature selection using ranking methods and classification algorithms." Yugoslav Journal of Operations Research. Vol.1, 2016

[26]. Ambusaidi MA, He X, Nanda P, Tan Z. Building an intrusion detection system using a filter-based feature selection algorithm. IEEE transactions on computers.vol 1, pp.2986-98,2016

[27]. Chung AG, Shafiee MJ, Wong "A. Random feature maps via a Layered Random Projection (LARP) framework for object classification"., InImage Processing (ICIP), pp. 246-250, IEEE, 2016

[28]. Muda Z, YassiClustering and Naive Bayes Classification for Intrusion Detection", Journal of IT in Asia. Vol 4, issue 1, pp: 13-25, 2016.