ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Analysis of Machine Learning Techniques for Intrusion Detection

Anjum Khan [1], Anjana Nigam [2]

P.G. Student, Department of Computer Science & Engineering, SIRT, Bhopal, India[1]

Professor, Department of Computer Science & Engineering, SIRT, Bhopal, India[2]

ABSTRACT: As network applications grow rapidly, network security mechanisms require more attention to improve speed and accuracy. The evolving nature of new types of intrusion poses a serious threat to network security: although many network security tools have been developed, the rapid growth of intrusive activities is still a serious problem. Intrusion detection systems (IDS) are used to detect intrusive network activity. Machine learning and data mining techniques have been widely used in recent years to improve intrusion detection in networks. These techniques allow the automatic detection of network traffic anomalies. One of the main problems encountered by researchers is the lack of data published for research purposes. In this research work the proposed model for intrusion detection is based on particle swarm optimization and knn classifier. The work is performed in divided into two stages. In first stage analysis of proposed machine learning approaches is performed on existing database i.e. NSL-KDD database and UNSW_NB15. In second stage generation of database on real packets extraction from network and performance analysis of different machine learning approaches (k-Nearest Neighbour, Random Forest, Support Vector Machine, Naïve Bayes and AdaBoost) on this database. In both stages feature reduction technique is applied on the dataset so that efficient result can be obtained in selected features. As compared to some existing work the proposed algorithm proves its efficiency in terms of high accuracy and low false alarm rate.

KEYWORDS: Intrusion Detection, Machine Learning, Classification, Accuracy, FPR, FNR, FAR.

I. INTRODUCTION

Intrusion Detection System (IDS)s are security tools that detect intrusions to a network or a host computer. An IDS is either host based or network based. A host based IDS detects attacks on a host computer, whereas, a network based IDS, also called Network Intrusion Detection System (NIDS), detects intrusions into a network by analyzing network traffic and are generally installed in network gateway or server. Host based intrusion detection systems can be divided into four types, namely (a) File System Monitors, (b) Log file analyzers, (c) Connection analyzers, (d) Kernel-based IDS [1, 2].
Furthermore, based on the data analyzing technique there are principally 2 classes of IDSs, signature-based and anomaly primarily based. A signature-based system detects attacks by analyzing network data for attack signatures hold on in its database. this kind of IDS detects previously best-known attacks, whose signatures are stored in its database. On the other hand, an anomaly-based IDS appearance for deviations from traditional behavior of the subject. Anomaly-based systems are capable of detecting novel attacks [3-7].
Here some very common methods given which are used by intruders to gain control of computers are Trojan horse, Back door, Denial of Service, Email-borne Viruses, Packet sniffing, Spoofing, etc. It is clear from the study that a network packet has 42 features and the four simulated attacks such as [8-12]:
Denial of Service (DoS) attack: Over usage of the bandwidth or non availability of the system resources leads to the DoS attacks. Examples: Teardrop and Smurf.
User to Root (U2R) Attack: Initially attacker access normal user account, later gain access to the root by exploiting the vulnerabilities of the system. Examples: Perl, Load Module and Eject attacks.
Probe Attack: Have an access to entire network information before introducing an attack. Examples: ipsweep, nmap attacks.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Root to Local (R2L) Attack: By exploiting some of the vulnerabilities of the network attacker gains local access by sending packets on a remote machine.
Machine learning techniques can be effective for detecting intrusions. Many Intrusion Detection Systems are modeled based on machine learning techniques [13,14,15]. Learning algorithms are designed either on offline dataset or real data collected from university or organizational networks.
Usually machine learning techniques is classified into 2 classes i.e. supervised Learning and unsupervised Learning. In supervised learning the training dataset is instantly accessible together with its target vector. The learner learns from available data taking guidance of the output vector [16,17,18].
In contrast to supervised learning, unsupervised learning systems learn from its atmosphere. Systems learn from coaching knowledge; however there's no target vector accessible. Some usually used machine learning techniques within the field of intrusion detection are like Artificial Neural Network (ANN), decision Tree, Support Vector Machine, Bayesian Classification, Self-organizing Map, etc [19-25].

II. RELATED WORK

Taeshik Shon [26] designed a framework consists of two Main components: Genetic algorithm (GA) for the characteristic selection machine and vector carrier (SVM) for the packet behavior classification.
YadigarImamverdiyev [27] discussed that intrusion detection systems are one of the most relevant security features against network attacks. Machine learning methods are used to analyze network traffic parameters in the presence of attack signs. This article discusses the extreme machine learning method for detecting intrusions in network traffic. The experimental results lead to the conclusion of the practical significance of the proposed approach to detect attacks in network traffic.
Athanasios Tsiligkaridis [28] developed a method to detect atypical bottlenecks in traffic City of Boston Our motivation is to detect these traffic jams which are often caused by an event (for example, an accident, a lane closure, etc.) and allow the city to intervene before congestion roads and adjacent roads are negatively affected. Using a data set on the traffic jams provided by the city of Boston presents a new detection system for the identification of anomalous jams. We demonstrate its effectiveness by using it to identify traffic jams that cannot be explained with typical traffic patterns.
Bhanu Vrat et al. [29] discussed that detection of anomalies is important requirement to protect a network against the strikers. Detects attacks on a network the analysis of the behavioral model was a important field of study for many researchers application systems in IPv4 and IPv6 networks. For accurate detection of anomalies, it is essential implement and use effective data mining methodology such as machine learning. In this article we considered a model of anomaly detection that uses machine learning algorithms for data mining in a network to detect anomalies present at any time. This the proposed model is evaluated against denial of service Attacks (DOS) in IPv4 and IPv6 networks selecting the most common and obvious features of IPv6 and IPv4 networks to optimize detection. The results also show that the proposed system can detects most IPv4 and IPv6 attacks effectively way.
Shuai Zhao et al. [30] presented a new structure for the real Detection of network traffic anomalies through machine learning algorithms. The proposed prototype system uses large existing ones data processing framework like Apache Hadoop, Apache Kafka and Apache Storm in conjunction with machine learning techniques and tools. Our approach is a system of Real-time processing and network flow analysis in real time the data collected on the University's university network from Missouri-Kansas City. Furthermore, the network anomaly the models have been identified and evaluated using machine learning techniques We present preliminary results on the detection of anomalies with data from the campus network.
Khadija Hanifi et al. [31] discussed that network attacks are exceptional cases they are not observed in the normal behavior of the traffic. In this work, to detect network attacks, using the k-means algorithm a new semi-supervised anomaly detection system was designed and implemented. During the training phase, normal samples were split into clusters by applying the k-means algorithm. So in To be able to distinguish between normal and abnormal samples, based on their distance from cluster centers and using a validation data set, a threshold value has been calculated. New samples that are far from cluster centers more than the threshold value is detected as anomalies. We used NSL-KDD- a data set labeled network connection traces - to test ours the effectiveness of the method. The experiments result in NSL-KDD dataset, shows that we have reached an accuracy of 80.119%.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

WathiqLaftah Al-Yaseen et al. [35] proposes a multi-level hybrid intrusion detection model that uses support vector machine and extreme learning machine to improve the efficiency of detecting known and unknown attacks. A modified K-means algorithm is also proposed to build a high-quality training dataset that contributes significantly to improving the performance of classifiers. The modified K-means is used to build new small training datasets representing the entire original training dataset, significantly reduce the training time of classifiers, and improve the performance of intrusion detection system. The popular KDD Cup 1999 dataset is used to evaluate the proposed model. Compared with other methods based on the same dataset, the proposed model shows high efficiency in attack detection, and its accuracy (95.75%) is the best performance thus far.
Hebatallah Mostafa et al. [36] presents a framework for selecting features for effective detection of network anomalies using a different machine learn classifiers the framework applies different strategies using filter and wrapper selection methods. The purpose of this framework is to select the minimum number of features that achieve maximum precision. UNSW-NB15 data set is used in the experimental results to evaluate the frame. The results show that using 18 features of a filter the classification methods and apply J48 as a classifier, an accuracy of 88% is achieved.

III. METHODOLOGY

The biggest challenge for today is to protect the users from Intrusion due to wide use of internet. Intrusion Detection Systems (IDS) are one of the security tools available to detect possible intrusions in a Network or in a Host.
Research showed that application of machine learning techniques in intrusion detection could achieve high accuracy rate as well as low false alarm rate. Accurate predictive models can be built for large data sets using supervised machine learning techniques, that is not possible by traditional methods.
IDS learns the patterns by the training data, so it can detect only the known attack, new attacks cannot be identified. This research work is based on designing an optimized feature based classifier and performing analysis on three different datasets.
This section describes the proposed hybrid model for intrusion detection. The NSL-KDD dataset, UNSW_NB15 as well as real dataset can be used as a benchmark to evaluate the performance of the proposed model. The algorithm flow of the proposed method is described as follows:
Following steps will be used to build the proposed model for intrusion detection:
Step 1: Convert the symbolic attributes protocol, service, and flag to numerical.
Step 2: Normalize data to [0,1].
Step 3: Separate the instances of dataset into two categories: Normal, Attack.
Step 4: Feature Reduction and Extraction.
Step 5: The data set is divided as training data and testing data.
Step 5: Train classifier with these new training datasets.
Step 6: Test model with dataset.
Step 7: Finally computing and comparing Accuracy and FAR for different classifiers.

A. Proposed Algorithm

procedure Intrusion_Behaviour(dataset)
Input: D {dataset};
Output: Label {Normal, Intrusion};
Step1: For each instance in D, do
Find feature vector (V)
Step 2: For each V do
Feature Reduction using Particle Swarm Optimization
Step 3: Data classification using Classifiers such as k-Nearest Neighbour, AdaBoost, Random Forest and Support Vector Machine.
Step 4: Determine the total class label
Find
True_positive (TP)
True_negative (TN)
False_positive (FP)

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

False_negative (FN)
Step 5: Find Performance Parameters
Step 6: Predict Intrusion behavior as
if (class=1) Test_Data = Intrusion
else_if(class =0) Test_Data=Normal
end for

B. Proposed Methodology

The proposed algorithm flow diagram of intrusion detection model is illustrated in figure 1. The proposed framework consists of three phases i.e. Preprocessing, Post Processing Phase and Intrusion Detection Phase. Below each stage is described individually in details.

Preprocessing

This stage purpose is to preprocess the database file in which there is conversion of symbolic attributes protocol, service, and flag in numerical is done. Further data is normalized.

Post-Processing Phase

Once pre-processing is applied, the pre-processing Module creates the Feature Vector matrix of dataset that represents in which each row i represents the instances and j represents the packet attributes.

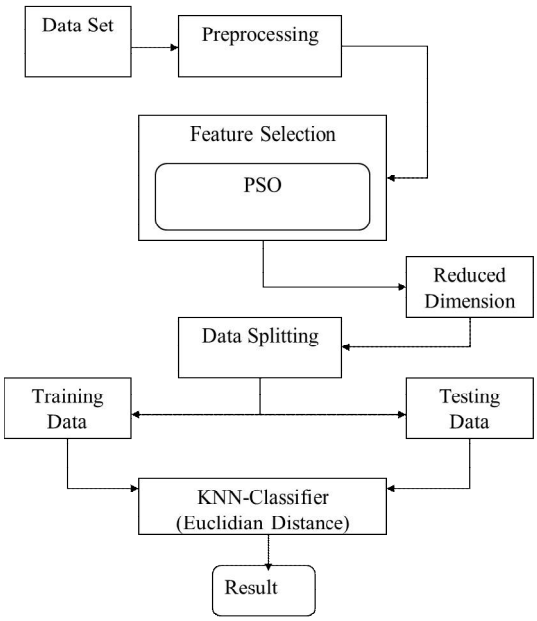

Figure 1: Proposed Flow Diagram of Intrusion Detection System

Feature Selection

The aim Feature selection phase is to further select only those features from the database which are relevant for proper classification of the dataset and consequently reduces the feature space dimension so as to reduce complexity by removing irrelevant data. This task is accomplished by using the Particle Swarm Optimization (PSO).

PSO Algorithm

The basic process of the PSO algorithm is given by:

Step 1: (Initialization) Randomly generate initial particles. For the PSO algorithm, the complete set of features is represented by a string of length N.

Step 2: (Fitness) Measure the fitness of each particle in the population. The selection of this fitness function is a crucial point in using the PSO algorithm, which determines what a PSO should optimize. Here, the task of the PSO algorithm is

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

to find the global minimum value according to the definition of the fitness function. The definition of the fitness function for the basic method is simply the accuracy of detection.
Step 3: (Update) Compute the velocity of each particle.
Step 4: (Construction) For each particle, move to the next position.
Step 5: (Termination) Stop the algorithm if the termination criterion is satisfied; return to Step 2 otherwise.

Classification
This section introduces an innovative computational intelligence framework for the purpose of analyzing malicious data in the context of intrusion detection. Let V(n*k) be the reduced database by dimensions matrix containing only the selected features which resulted after applying PSO. This module takes as input matrix V(n*k) and applies a KNN classifier as a learning algorithm. The learning algorithm works as follows:
K-nearest neighbor (KNN) classification algorithm is a data mining algorithm which is theoretically mature with low complexity.
The basic idea is that, in a sample space, if most of its *K* nearest neighbor samples belong to a category, then the sample belongs to the same category.
The nearest neighbor refers to the single or multidimensional feature vector that is used to describe the sample on the closest, and the closest criteria can be the Euclidean distance of the feature vector.
Let m be the number of training data samples. Let p be an unknown point.
Store the training samples in an array of data points arr[]. This means each element of this array represents a tuple (x, y).
for i=0 to m: Calculate Euclidean distance d(arr[i], p).
Make set S of K smallest distances obtained. Each of these distances correspond to an already classified data point.
Return the majority label among S.

IV. SIMULATION RESULTS

A. Description of Dataset

NSL-KDD Dataset
The inherent drawbacks in the KDD cup 99 dataset [9] has been revealed by various statistical analyses has affected the detection accuracy of many IDS modelled by researchers. NSL-KDD data set [3] is a refined version of its predecessor. It contains essential records of the complete KDD data set. There are 42 attributes in NSL-KDD Dataset. The 42nd attribute contains data about the various 5 classes of network connection vectors and they are categorized as one normal class and four attack class. The 4 attack classes are further grouped as DoS, Probe, R2L and U2R. The description of the attack classes.
UNSW-NB 15 Dataset
The raw network packets of the UNSW-NB 15 data set was created by the IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) for generating a hybrid of real modern normal activities and synthetic contemporary attack behaviours.
Tcpdump tool is utilised to capture 100 GB of the raw traffic (e.g., Pcap files). This data set has nine types of attacks, namely, Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode and Worms. The Argus, Bro-IDS tools are used and twelve algorithms are developed to generate totally 49 features with the class label.

B. Performance Evaluation Measures

To evaluate the proposed algorithm, it is concentrated on three indications of performance: detection rate, accuracy and False Alarm Rate (FAR).
If one sample is an anomaly and the predicted label also stands anomaly, then it is called as true positive (TP).
If one sample is an anomaly, but the predicted label stands normal, then it is called as false negative (FN).
If one sample is a normal and the predicted label also stands normal, then it is true negative (TN).
If one sample is normal, but the predicted label stands anomaly, then it is termed as false positive (FP).
TP stands the number of true positive samples, FN stands the number of false negative samples, FP stands the number of false positive samples, and TN stands the number of true negatives.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

From equation (1) and (4), the accuracy, False Positive rate (FPR), False Negative Rate (FNR) and False Alarm rate (FAR) is calculated.

Accuracy = (TP+TN)/(TP+TN+FP+FN) (1)

False Negative Rate (FNR) = FN/(FN+TP) (2)

False Positive Rate (FPR) =FP/(FP+TN) (3)

False Alarm Rate (FAR)= (FPR+FNR)/2 (4)

C. Result Analysis

For performance evaluation, this research work is divided into three sections, where three different datasets are analyzed using different classifiers. The performance evaluation are performed using feature reduction technique i.e. PSO. By applying PSO algorithm feature reduction technique over NSL-KDD dataset it has been observed that best result is obtained at 30 features out of 42 features. Whereas after applying PSO algorithm feature reduction technique over UNSW-NB 15 dataset it has been observed that best result is obtained at 35 features out of 49 features. Table I shows the performance evaluation of PSO-KNN algorithm over three datasets.

Table I: Performance Evaluation of Proposed Algorithm

DATASETS	ACCURACY	FNR	FPR	FAR
NSL-KDD (30 features)	98.5755	1.1514	1.6642	1.4078
UNSW_NB15 (35 Features)	97.9796	0.001	2.0518	1.0259
Real Data (11 Features)	86.4593	12.9398	16.0377	14.4888

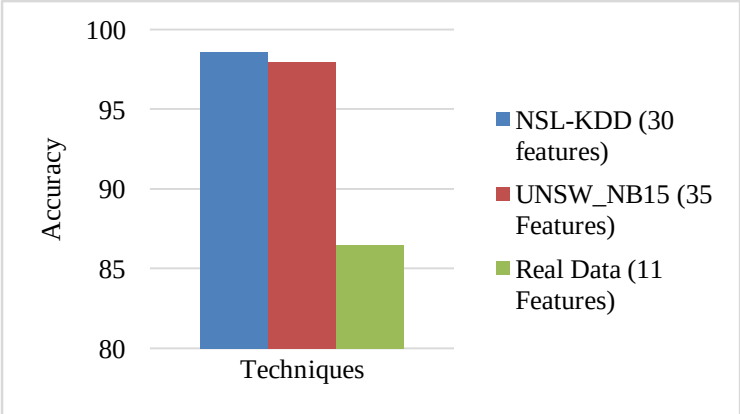

Fig. 2: Result Analysis of Proposed Algorithm for Accuracy

Figure 2 illustrates the performance of the PSO-KNN algorithm and it is concluded that the result analysis of for NSL-KDD is about 98.5% and UNSW-NB 15 is about 98% and for real dataset is about 86%.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

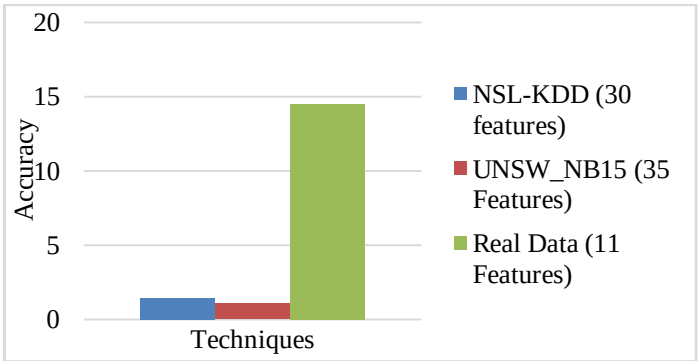

Fig. 3: Result Analysis of Proposed Algorithm for FAR

Figure 3 illustrates the performance of the PSO-KNN algorithm and it is concluded that the result analysis of for NSL-KDD is about 98.5% and UNSW-NB 15 is about 98% and for real dataset is about 86%. Three dataset comparative results are illustrated in below sections:

Comparative Result Analysis of NSL-KDD Dataset

This comparative result analysis is performed by comparing the result analysis of proposed work with existing work of [35]

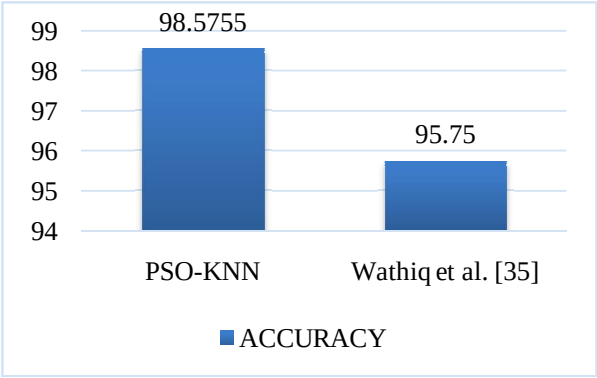

Fig. 4: Accuracy Comparative Analysis of Proposed Algorithm using NSL-KDD Dataset

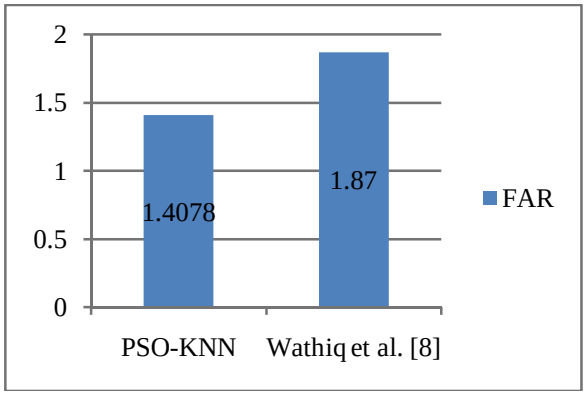

Fig. 5: FAR Comparative Analysis of Proposed Algorithm using NSL-KDD Dataset

Figure 4and 5 illustrates the performance of the PSO-KNN algorithm and it is concluded that the result analysis of proposed work with existing work of [35].

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Comparative Result Analysis of UNSW-NB 15 Dataset

This comparative result analysis is performed by comparing the result analysis of proposed work with existing work of [37].

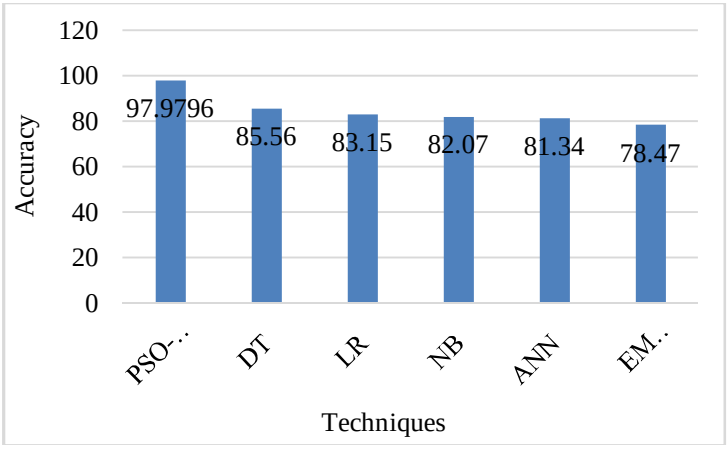

Fig. 6: Accuracy Comparative Analysis of Proposed Algorithm using UNSW-NB 15 Dataset

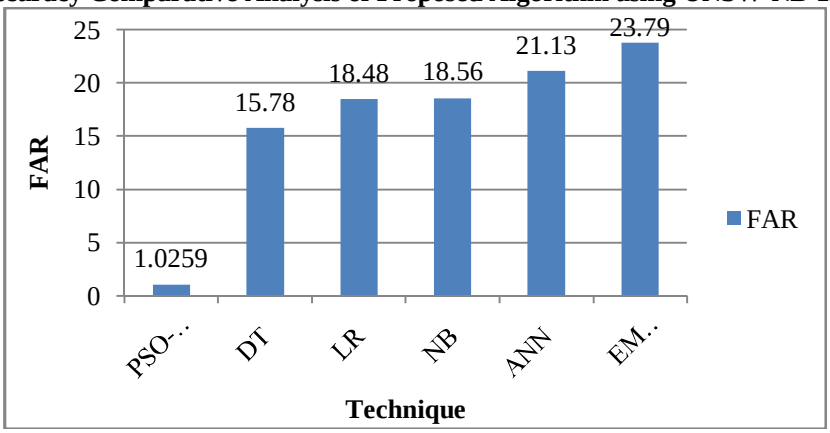

Fig. 7: Accuracy Comparative Analysis of Proposed Algorithm using UNSW-NB 15 Dataset

Figure 6 and 7 illustrates the performance of the PSO-KNN algorithm and it is concluded that the result analysis of proposed work with existing work of [37].

Result Analysis of Real Dataset

Presented experiment showing day-wise results. Total 5 days results are presented by capturing live packets from SIRT lab started at approx. 12:00 PM and Stopped at approx. 4:00 PM every day. The dataset was prepared by designing rules as illustrated in NSL-KDD dataset and observed 5 different types of data packets i.e. DOS, U2R R2L, probe and normal data packets. Each packet are analyzed and their dataset was prepared. During capturing following analysis is performed as shown above in Table II.

Table II: Real Time Analyzed Packets

Days	Total Received Packets				
	DOS	Probe	U2R	R2L	Normal
Day 1	423	200	143	165	532
Day 2	200	174	157	194	256
Day 3	134	165	154	102	431
Day 4	89	176	138	121	369
Day 5	146	149	251	64	534

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

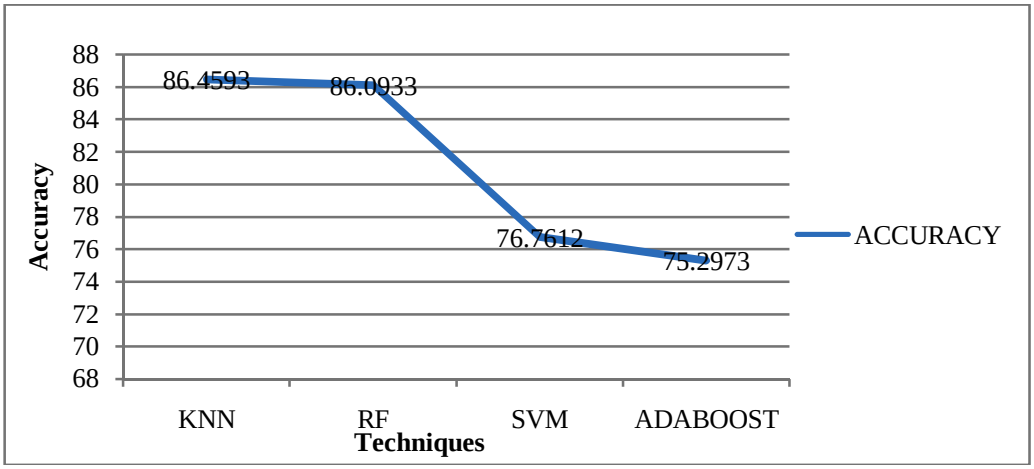

Fig. 8: Accuracy Analysis of Proposed Algorithm using Real Dataset

Figure 8 and 9 illustrates the performance of the PSO-KNN algorithm and it is concluded that the result analysis of proposed work with real dataset.

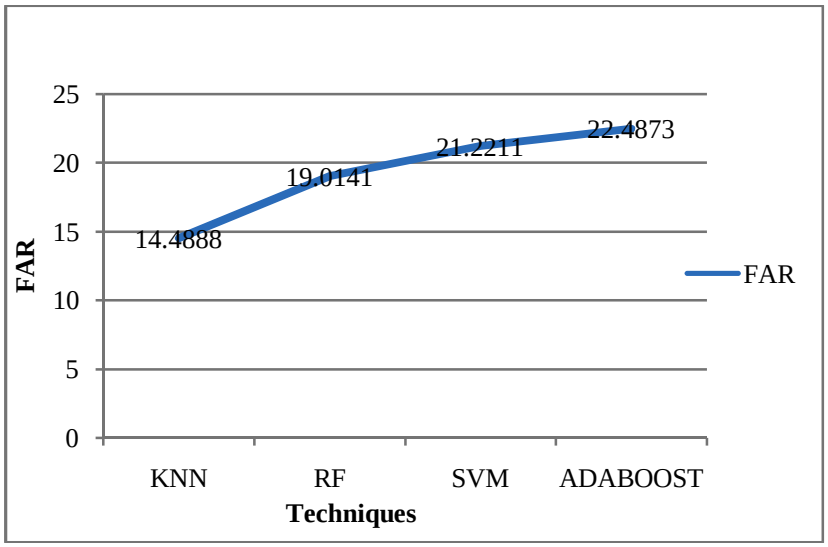

Fig. 9: FAR Analysis of Proposed Algorithm using Real Dataset

V. CONCLUSION

In this research work the proposed model for intrusion detection is based on particle swarm optimization and knn classifier. The work is performed in divided into two stages. In first stage analysis of proposed machine learning approaches is performed on existing database i.e. NSL-KDD database and UNSW_NB15. In second stage generation of database on real packets extraction from network and performance analysis of different machine learning approaches (k-Nearest Neighbour, Random Forest, Support Vector Machine, Naïve Bayes and AdaBoost) on this database. In both stages feature reduction technique is applied on the dataset so that efficient result can be obtained in selected features. The result analysis performed on three datasets using PSO-KNN algorithm concludes that:
NSL-KDD Dataset gives accuracy of about 98.57% and FAR of about 1.40%.
UNSW-NB15 Dataset gives accuracy of about 97.91% and FAR of about 1.02%.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

Real Dataset gives accuracy of about 86.45% and FAR of about 14.48.%..

REFERENCES

1. Garcia-Teodoro, P., "Anomaly-based network intrusion detection: techniques", systems and challenges. Comput. Security vol. 28. issue, pp. 18–28, 2009.
2. Sufyan T Faraj Al-Janabi, Hadeel Amjed Saeed, "A neural network-based anomaly intrusion detection system", IEEE, 2011.
3. J. Ryan, M. Lin, and R. Miikkulainen, "Intrusion Detection with Neural Networks," Conference in Neural Information Processing Systems, 943–949.
4. A. K. Ghosh and A. Schwartzbard, "A Study in Using Neural Networks for Anomaly and Misuse Detection," Conference on USENIX Security Symposium, Volume 8, pp. 12–12, 1999.
5. P. L. Nur, A. N. Zincir-heywood, and M. I. Heywood, "Host-Based Intrusion Detection Using Self-Organizing Maps," in Proceedings of the IEEE International Joint Conference on Neural Networks, pp. 1714–1719, 2002.
6. K. Labib and R. Vemuri, "NSOM: A Real-Time Network-Based Intrusion Detection System Using Self-Organizing Maps," 2000.
7. Sharma, R.K., Kalita, H.K., Issac, B., "Different firewall techniques: a survey", International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2014.
8. Meng, Y.-X., "The practice on using machine learning for network anomaly intrusion detection", International Conference on Machine Learning and Cybernetics (ICMLC), vol. 2, IEEE, 2011.
9. SumaiyaThaseen Ikram, Aswani Kumar Cherukuri, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM", Journal of King Saud University –Computer and Information Sciences, 2016.
10. Manjula C. Belavagi and BalachandraMuniyal, "Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection, Procedia Computer Science", Elsevier, 2016.
11. Saad Mohamed Ali Mohamed Gadal and Rania A. Mokhtar, "Anomaly Detection Approach using Hybrid Algorithm of Data Mining Technique", International Conference on Communication, Control, Computing and Electronics Engineering, IEEE, 2017.
12. Ibrahim, H. E., Badr, S. M., &Shaheen, M. A., "Adaptive layered approach using machine learning techniques with gain ratio for intrusion detection systems", International Journal of Computer Applications, vol. 56, issue 7, pp. 10–16, 2012.
13. Wen Feng, Qinglei Zhang, Gongzhu Hu, Jimmy Xiang Huang, "Mining network data for intrusion detection through combining SVMs with ant colony networks", Elsevier, Vol 37, pp 127-140, 2014.
14. Shi-JinnHorng, Ming-Yang Su, Yuan-Hsin Chen, Tzong-Wann Kao, Rong-Jian Chen, Jui-Lin Lai, Citra Dwi Perkasa, "A novel intrusion detection system based on hierarchical clustering and support vector machines" Expert Systems with Applications, Elsevier, vol. 38, pp. 306–313, 2011.
15. O.Y.Al-Jarrah, O. Alhussein, P.D.Yoo, S. Muhaidat, K.Taha and K. Kim, " Data Randomization and Cluster-based Partitioning for botnet intrusion detection", IEEE Transactions on Cybernetics, vol. 46, no. 8, pp. 1796-1806, 2016.
16. Solane Duque, Dr. Mohd. Nizam Bin Omar, "Using Data Mining Algorithm for Developing a Model for Intrusion Detection System(IDS)", procedia Computer Science 61 (2015) 46-51.
17. Paul Dokas, LeventErtoz, Vipin Kumar, Aleksandar Lazarevic, Jaideep Srivastava, Pang-Ning Tan , " Data Mining for Network Intrusion Detection ", University of Minnesota, Minneapolis, MN 55455, USA.
18. Wenke Lee, Savatore J. Stolfo, Kui W. Mok, " A Data Mining Framework for Building Intrusion Detection Models", Computer Science Department , Columbia University 500 west 120th street, New York 10027.
19. Mathew G. Schultz and Eleazar Eskin and Erez Zadok, "Data Mining Methods for Detection Of New Malicious Executables", Department of Computer Science Columbia University.
20. Aasia Abdullah and KhaledaAfroaz," Data Mining Approaches on Network Data: Intrusion Detection System", International Journal of Advanced Research in Computer Science Volume 8, No. 1, Jan-Feb 2017.
21. Ashok Chalak, Naresh D Harale and Rohini Bhosale, "Data Mining Techniques for Intrusion Detection and Prevention", IJCSNS International Journal of Computer Science and Network Security, Vol. 11 No. 8, August 2011.
22. G. V. Nadiammai and M. Hemalatha, "Effective Approach Towards Intrusion Detection System Using Data Mining Techniques", Department of Computer Science, Karpagam University, Coimbatore 641021, Tamilnadu, India.
23. Nutan Farah Haq, MusharratRafni, Abdur Rahman Onik, Faisal Muhammad Shah, Md. Avishek Khan Hridoy and Dewan Md. Farid, " Application of machine Learning Approaches in Intrusion Detection System : A Survey", (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No. 3, 2015.
24. Kuang, F., Xu, W., & Zhang, S., "A novel hybrid KPCA and SVM with GA model for intrusion detection", Applied Soft Computing Journal, vol. 18, pp. 178–184, 2014.
25. Prasanta Gogoi, D.K. Bhattacharyya, B. Borah1 and Juga, K. Kalita, "MLH-IDS: A Multi-Level Hybrid Intrusion Detection Method", The Computer Journal, Vol. 57 issue 4, pp. 602-623, 2014.
26. Taeshik Shon "A Machine Learning Framework for Network Anomaly Detection using SVM and GA", IEEE, 2005.
27. YadigarImamverdiyev "Anomaly detection in network traffic using extreme learning machine", IEEE, 2016.
28. Athanasios Tsiligkaridis "Anomaly Detection In Transportation Networks Using Machine Learning Techniques", IEEE, 2017.
29. Bhanu Vrat et al "Anomaly Detection in IPv4 and IPv6 Networks Using Machine Learning", IEEE, 2015.
30. Shuai Zhao et al "Real-Time Network Anomaly Detection System Using Machine Learning", IEEE, 2015.
31. Khadija Hanifive Hasan Bank "Network Intrusion Detection Using Machine Learning Anomaly Detection Algorithms" , IEEE, 2016.
32. He, L., "An improved intrusion detection based on neural network and fuzzy algorithm. Journal of Networks, vol. 9, issue 5, pp. 1274–1280, 2014.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 5, May 2018

33. Hoque, M. S., Mukit, M. A. &Bikas, M. A. N., "An implementation of intrusion detection system using genetic algorithm", International Journal of Network Security & Its Applications, vol 4, issue 2, pp. 109–120, 2012.
34. Nour Moustafa& Jill Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set", Information Security Journal: A Global Perspective, 2016.
35. WathiqLaftah Al-Yaseen, Zulaiha Ali Othman, MohdZakree Ahmad Nazri, "Multi-Level Hybrid Support Vector Machine and Extreme Learning Machine Based on Modified K-means for Intrusion Detection System", International Journal in Expert Systems With Applications, Elsevier, 2017.
36. Hebatallah Mostafa Anwer et al., "A Framework for Efficient Network Anomaly Intrusion Detection with Features Selection", IEEE, 2018.
37. Nour Moustafa& Jill Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set", Information Security Journal: A Global Perspective, 2015.