ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Network Intrusion Detection System (NIDS) using Machine Learning Perspective

Sagar Dhende [1], R. B. Ingle[2]

M.E Scholar, Dept of Computer Engineering, PICT Pune, SPPU, Pune, India[1]

Professor, Dept of Computer Engineering, PICT Pune, SPPU, Pune, India[2]

ABSTRACT: Many intrusion detection system are rule based which can not detect novel attacks. Moreover, rule based technique is time consuming due to the encoded rule manually and it highly depend on the prior knowledge of the known attacks. Therefore, we proposed network based intrusion detection system (NIDS) using machine learning technique. NIDS is meant to be a device or a system application that monitor a network traffic and event occurring in a computer system. In network security intrusion detection system play a major role to detect different kinds of attacks. The machine learning technique can be used to increase the attack detection performance. In this paper Network intrusion detection system is proposed with the method of principle component analysis (PCA) and support vector machine (SVM). This proposed method was tested on KDDCup dataset and attack detection accuracy is compared to decision tree and naive bayes algorithms.

KEYWORDS:Intrusion Detection, Classification, Machine Learning, User Behavior.

I. INTRODUCTION

Tremendous growth of internet usage in daily life raise the concern about how to secure network from different kinds of possible attacks. The existing solution is incapable to fully protecting network and internet application against the threats and vulnerabilities from cyber attacks. So the network security is more important than the traditional network security. There are various security technologies are available to protect the network based system but there are still many undetected threats. Intrusion detection system is highly recommended to provide security of computer network. In general intrusion detection system classifying into two types one is Host Based Intrusion Detection System (HIDS) and other one is Network Based Intrusion Detection System (NIDS).

Host Based Intrusion Detection System (HIDS) is capable to analyzing and monitoring computer system or network packet on network. It work similar to the network based intrusion detection system but difference in HIDS and NIDS is the HIDS is only install on certain intersection point such as routers , servers while NIDS is install on every host machine. Our research approach is to present Network based intrusion detection system to detect various attacks and unknown behaviors using machine learning approach. Network based intrusion detection system contain two techniques one is rule based detection technique and other is anomaly based detection technique. Rule based technique is used to discovers known attacks, while anomaly detection technique is used to detect the known and unknown attacks. In the rule based technique traffic of network is very huge so that the process of the manually encoded rules is very slow and expensive. Cyber security professionals and experts are needed to change or modify the rules manually by using the rule driven languages. To overcome such kinds of limitations of rule based technique, anomaly based detection technique is used.

In this work we proposed network based intrusion detection system using machine learning approach. There are two kind of machine learning technique: supervised algorithm and unsupervised algorithm. The unsupervised algorithm take a unlabeled data or test data as input and detect the malicious activities based on the assumptions, but is not applicable to many cases such as DoS attacks, outage or misconfiguration the supervised algorithm take a labeled data as its input and it can detect wide range of malicious activities and anomalies which is unsupervised cannot. The

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

contribution of this work are finding out the different kinds of threats, vulnerabilities or cyber attacks using two approaches of machine learning such as Principal component analysis (PCA) and Support vector machine (SVM)

II. LITERATURE SURVEY

1. Mohammed Ambusaidi and Priyadarshi Nanda. (2016) [1] described flexible mutual information feature selection technique which is useful to select optimal features in the process of classification. This technique reduce the computational cost and increase the accuracy.
2. Jiong Zhang et al. (2008) [2] random forest algorithm is apply in three techniques such as anomaly technique, misuse technique, and hybrid detection technique. in this work the hybrid detection method that increases the high performance as it combining the benefits or characteristics of both technique anomaly detection technique and misuse detection technique. the main advantage of this work is it build patterns automatically by avoiding the rule manually.
3. Fang-Yie Leu et al. (2017) [3] This system is used to detect insider vulnerabilities, attacks, threats at system call side, In this work forensic and data mining techniques are used. The main advantage of this work is prevent the system from the insider threats, malicious activities and attacks effectively.
4. Kriangkrai Limthong and Thidarat Tawsook (2012) [4] conducted an experiments to study and analysis of relationship between interval based features of network traffic by using two approaches of machine learning: k-nearest neighbor and naive Bayes. In this work they try to help network administrator and researchers to select proper algorithm of machine learning for their work.
5. S. Yeldi, S. Gupta, R Ingle et al. (2009) [5] describes the intrusion detection system by using the honeypot, This system mainly focus on the honeypots possibilities in an educational environment also focus on productive environment. The honeypot is very useful to attract and divert to attacker from their target or goals by emulating real service.
6. Shuai Zhao, Deep Medhi, Mayanka Chandrashekhar, et al. (2017) [7] presented novel based framework for the real time malicious activities, anomaly detection or vulnerabilities using the machine learning techniques, along with the machine learning technique the novel framework is also used big data frameworks namely Apache storm, Apache Hadoop, Apache Kafka etc.
7. Nguyen Thanh Van,Tran Ngoc Thinh, Le Thanh Sach. (2017) [8] presented a deep learning to implement an anomaly based network intrusion detection system. In these technique two deep learning techniques such as Stacked Autoencoder and Stacked RBM is used to detect anomaly with high accuracy and attack is also classified into four group to increase accuracy.
8. Nabila Farnaaz and M.A.Jabbar (2017) [9] In this random forest algorithm is used to built a model of network intrusion detection system. random forest is ensemble classifier is used for classification and perform well as compare to other classifier techniques for effective classification of attacks.

III.RELATED WORK

1. Data Set

The KDD cup'99 dataset used in this experiment, dataset 10 percent KDD Cup'99 is used as training dataset and KDD Cup'99 is used as testing data set to detect attack occurred in the network. The whole data set consist of four types of attacks.

- Denial of Service(DOS): This is type of cyber attack in which attacker attack on network and try to making network resource inaccessible to its intended users.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

- Probing: Probing is the attack in which attacker scan networking device or machine to find out the machine's vulnerabilities or weakness.
- Root to Local (R2L): Attacker trying to gain access of user authentication of users machine, as attacker does not have any information or account on user machine.
- User to Root (U2R): This attack aim to acquire access permission from normal user to gain super user benefits.

Table 1. Classification of attacks

Attack Type	class	Attack Type	class
nmap	probing	xlock	R2L
Saint		warezmaster	
mscan		snmpgetattack	
Ipsweep		named	
portsweep		Imap	
satan		guess_passwd	
neptune	Normal	ftp_write	
normal		sendmail	
back		multihop	
mailbomb		phf	
pod		snmpguess	
smurf		worm	
udpstorm		httptunnel	
teardrop		buffer_overflow	U2R
processtable		perl	
land		rootkit	
apache2		xterm	
spy	R2L	sqlattack	
xsnoop		ps	
wazerclient		loadmodule	

2. Principle Component Analysis (PCA):

In network traffic there are various large amount of attributes which may be irrelevant to detect the attack type. hence the principle component analysis is used to remove that irrelevant attributes. Principle component analysis is the dimensionality reduction technique in which large dimension is reduced in small dimension, Due to the feature reduction the attack detection accuracy is increases and also improve the performance time.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

3. Support Vector Machine (SVM):

Support vector machine is the supervised learning that can analyze or perform pattern recognition and regression analysis task. Support vector machine is widely used for multi class classification by finding hyper-plane of training data set. The multi class support vector machine can be construct the multiple classes at the training time.

IV. PROPOSED METHODOLOGY AND DISCUSSION

The proposed system take input in the form of labeled and unlabeled data containing the threats and vulnerabilities. data per-processing is preprocessed over the training and test data and important features can classify or distinguish into class. Model for classification is trained using SVM classifier. Principle component analysis is used to attack recognition, where the intrusion detected on the test data and decision making is performed to take decision for intrusion detection or normal flow of data.

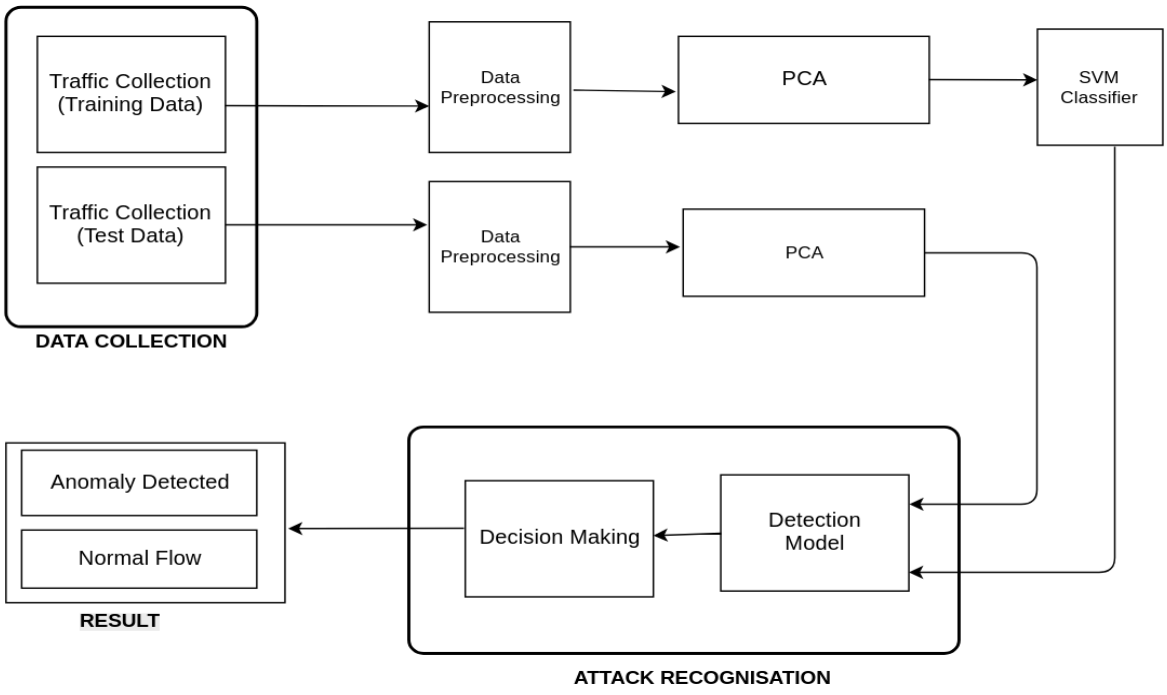

Fig 1: System architecture

V. EXPERIMENTAL RESULTS

In this experiment the accuracy of attack detection is increased by using the PCA and the SVM classifier. Table 2 shows the result table of 10 percent KDD Cup'99 dataset which is used as training dataset and KDD Cup'99 is used as testing data set. result of two data set is mentioned in the form of accuracy, precision, recall and f1-score. The attack detection accuracy of SVM is 91% which is listed in following table 2.

Table 2. Result Table

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Approaches	Accuracy	Precision	Recall	F1-Score
SVM	0.91	0.88	0.91	0.89
Decision tree	0.89	0.86	0.89	0.87
Naive Bayes	0.82	0.81	0.82	0.81

The figure 2 shows the performance analysis of NIDS where as comparison of three classification methods such as SVM, decision tree and naive bayes is shown with respect to precision, recall and f1-score. The accuracy of the SVM classifier is 91% which is more than the decision tree and naive bayes classifiers. The comparison of the SVM, decision tree and naive bayes is shown in following graph.

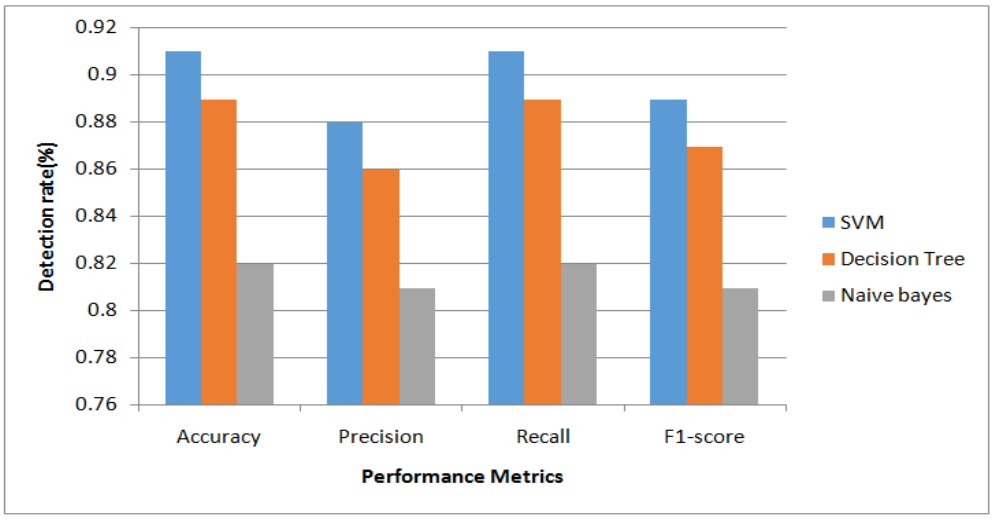

Fig 2: Performance Analysis

Standard metrics can be used to find out the precision, recall and f1-score, which is listed in table 3.

Metrics	Predicted	Predicted
	Normal	Attack
Normal	True negative (TN)	False Positive (FP)
Attack	False negatives (FN)	True positive (TP)

- Precision (P): Precision is the ratio of relevant prediction that are positive among the retrieved prediction.

$$P = TP / (TP+FP)$$

- Recall (R): Recall is the ratio of the relevant prediction from the retrieved prediction to the total relevant prediction.

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

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

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

R= TP / (TP+FN)

- F1- Score (F): F1-Score is the combination of precision and recall which maintain balance between precision and recall.

F = 2*Precision*Recall / Precision+Recall

VI. CONCLUSION

The Anomaly based technique complements the Signature based technique and helps in identifying the different novel attacks. The main objectives of the research is increasing the detection accuracy while avoiding the false positive alert. From the result it is observed that our proposed approach SVM gives more attack detection accuracy as compared to decision tree and naive bayes.

REFERENCES

[1] Mohammed a. Ambusaidi, Member, Priyadarsi Nanda and Zhiyun Tan, “Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm", IEEE transactions on computers, vol. 65,no. 10,october 2016.

[2] Jiong Zhang, Mohammad Zulkemine, and Anwar Haque, “Random-Forest-Based Network Intrusion Detection System", IEEE Transactions In Systems, Man, and Cybernetics, Part C (Applications and Reviews) (Volume: 38, Issue: 5, Sept. 2008).

[3] Fang-Yie Leu, Kun-Lin Tsai, Member, IEEE, Yi-Ting Hsiao, and Chao-Tung Yang, “An Internal Intrusion Detection and Protection System by Using Data Mining and Forensic Techniques" , IEEE systems journal, vol. 11, no. 2, june 2017.

[4] Kriangkrai Limthong and Thidarat Tawsook, “Network Traffic Anomaly Dtetection using Machine Learning Approach" , IEEE conference nov 2012.

[5] Sujata Yeldi, Sweta Gupta, Tanmay Ganacharya, Shirish Doshi, Dhanashree Bahirat, Prof, Rajesh Ingle, Anandamoy Roychowdhary, PICT, “Enhancing Network Intrusion Detection System with Honeypot", TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region.

[6] Juliette Dromard, Gilles Roudière, and Philippe Owezarski, “Online and Scalable Unsupervised Network Anomaly Detection Method", IEEE transactions on net work and service management, vol. 14, no. 1, march 2017.

[7] Shuai Zhao, Mayanka Chandrashekhar, Yugyung Lee, Deep Medhi, "Real-Time Network Anomaly detection System using Machine Learning", IEEE International Conference june 2015.

[8] Nguyen Thanh Van,Tran Ngoc Thinh, Le Thanh Sach, "An anomaly-based Network Intrusion Detection System using Deep learning", IEEE International Conference 2017.

[9] Nabila Farnaaz and M.A.Jabbar, "Random Forest Modeling for Network Intrusion Detection System", Nabila Farnaaz and M.A. Jabbar / Procedia Computer Science 89 (2016) 213 – 217.

[10] Ramana Rao Kompella, Student Member, IEEE, Sumeet Singh, and George Varghese, Member, IEEE, "On Scalable Attack Detection in the Network", IEEE/acm transactions feb 2007.

[11] Chee-Wooi Ten, Member, IEEE, Junho Hong, Student Member, IEEE, and Chen-Ching Liu, Fellow, IEEE, "Anomaly Detection for Cybersecurity of the Substations", ieee transactions on smart grid, vol. 2, no. 4, december 2011.

[12] Praneeth NSKH, Naveen Varma M, Roshan Ramakrishna Naik, “Principle Component Analysis based Intrusion Detection System Using Support Vector Machine”, IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016, India.

[13] Sumaiya Thaseen and Ch.Aswani Kumar, "Intrusion Detection Model Using fusion of PCA and optimized SVM" 2014 International Conference on Contemporary Computing and Informatics (IC3I), 978-1-4799-6629-5/14

[14] Sara Pourfallah, Amir H. Jafari et.al, "an intrusion detection algorithm for ami systems based on svm and pca", International Journal on Cybernetics & Informatics (IJCI)},Vol. 3, No. 4, August 2014.

[15] Soukaena Hassan Hashem, "efficiency of svm and pca to enhance intrusion detection system." in \textit{Journal of Asian Scientific Research,} 2013, 3(4):381-395.

[16] Gabriel Y. Keung, Member, IEEE, Bo Li, Fellow, IEEE, and Qian Zhang, Fellow, IEEE, “The Intrusion Detection in Mobile Sensor Network”,IEEE/ACM transactions on networking, vol. 20, no. 4, august 2012