International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Cyber Assault Detection in Network Using Machine Learning

[1] Yamini Mahitha Posina, [2]Anitha Padala, [3]Lohit Varma Vatsavai, [4]Srilatha Pediredla

[1,2,3,4]Department of CSE, Aditya Engineering College, Surampalem, A.P., India

ABSTRACT

Many researchers have been studying about intrusion detection systems when the cyber-attacks problem raised. The attacks are being attacked by the attackers through the internet. The network packets are sending through the client to server communication. It is very crucial to protect our data. Improving complexity of computer networks makes a challenging situation to classify the attack in network traffic. In this paper we are using nsl-kd99 data set for predicting the attack. Machine learning techniques gives more than 90 % accuracy for this NIDS detection. Advancements of machine learning techniques gives evaluation metrics and dataset collection. Intrusion detection system used to defend computer networks.

Keywords: NIDS, SVM, ANN, IDS, DOS, R2L, U2R, LDA, PCA.

INTRODUCTION

Nowadays the use of the internet is increasing day by day so cyber-attacks also take place through the internet. Due to these advancements, cybersecurity has gained attention. It leads to designing effective Network Intrusion detection systems (NIDS). Many IDS are developed for classifying the attacks. Many IDS methods have been proposed and each method has a different level of accuracy. Due to a large expansion of network size most of the applications are connected to the network nodes. A large amount of data is shared among the different network nodes. The security of these data becomes a challenging task for cyber security analysts. Existing Intrusion detection systems are divided into two types. They are misuse-based detection (it is also known as signature-based or knowledge-based detection) and anomaly-based detection (it is also known as behavior-based detection). Misuse-based detection system extracts the discriminative features and patterns from the attacks. It will also hand-code them into the system. This misuse-based detection system is very effective and efficient. Misuse-based detection systems must and should require updating the rules and signatures frequently. It is incapable to detect any novel or unknown attacks. An anomaly-based detection system can detect zero-day attacks. It can also adapt statistical methods, machine learning algorithms, and data mining algorithms to model the pattern of normal network behavior and detect anomalies as deviations from normal behavior. Over the years many researchers found machine learning-based and deep learning-based solutions to find these malicious attacks. The massive increase in network traffic and the resulting security threat has many challenges for NIDS systems to detect malicious attacks efficiently. Machine learning has the advantage of adaptability and capturing interdependencies. Machine learning-based techniques are their ability to learn and improve their performance over their time. The NSL-KD99 dataset is frequently used to detect cyber-attacks in machine learning. Classification algorithms such as Linear discriminant analysis (LDA) and Principal Component Analysis(PCA) to intrusion and classification anomaly. Machine learning emphasizes building a framework and enhances its execution based on previous results. The purpose of this paper is to give an overview of the advancement in ml techniques.

RELATED WORKS

In this model we can detect the attacks in the network and which type of attack is present in the network. Machine learning and deep learning approaches helps us to improve the complexity and efficiency to detect the cyber-attacks faster.

1. MACHINE LEARNING APPROACHES

Machine learning is the subset of Artificial Intelligence that can extract useful information from larger data sets. For the models, we can apply algorithms for our model. Machine learning algorithms such as Decision tree, Support vector machine(SVM), Artificial neural network.

1.1 DECISION TREE

Decision tree is one of the supervised machine learning algorithm. It is like a convolutional tree-like structure. They are two nodes decision node and a tree node in the decision tree and this algorithm also be used for both classification and regression problems. It will give possible solutions to decision-based conditions.

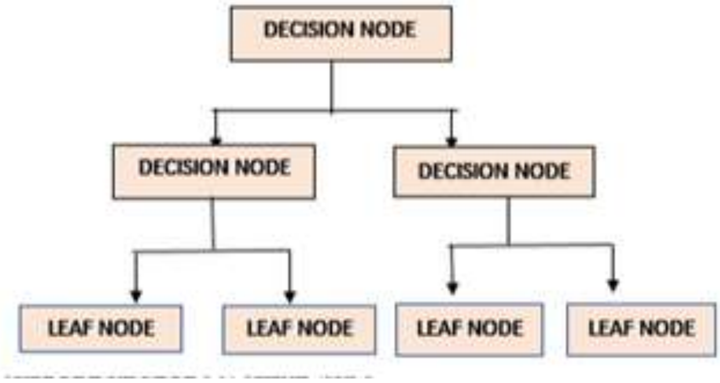


Fig 1. Decision tree

1.2 SUPPORT VECTOR MACHINE (SVM)

It is one of the popular supervised learning algorithms in machine learning. SVM finds a hyperplane in N-dimensional space and classifies data points. The dimension of the hyperplane depends upon features. For example, if the number of input features is two the hyperplane is a line and if it is three features then the hyperplane becomes a 2-D plane. It is used for linear and non-linear problems. Hyper-plane works as a decision boundary using a support vector machine. Non-linear problems use kernel functions. Network Intrusion Detection system the SVM algorithm improves its efficiency and accuracy.

1.3 ARTIFICIAL NEURAL NETWORK (ANN)

ANN is also a supervised learning algorithm. Artificial neural network contains artificial neurons which is also called units. It contains millions or dozens of units based upon the complexity of the system. ANN contains input layer, output layer and hidden layer. It has ability to perform non-linear modelling by learning from large datasets. Drawback is that the lower detection for lower attack classes. The number of neurons increases its accuracy.

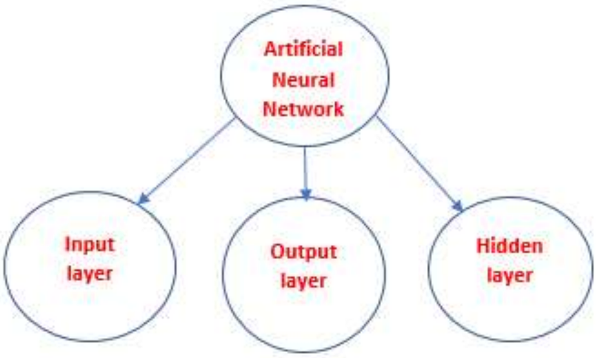
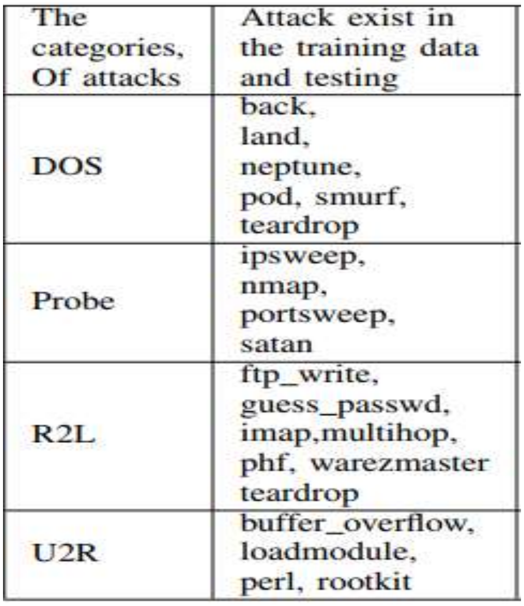


Fig 2. Artificial neural network

METHODOLOGY:

In this paper we are used Nsl-kd99 data sets to predict the attacks. In this Network Intrusion detection system application, we are predicting four types of attacks they are dos, probe, r2l, u2r.For predicting these attacks support vector machine algorithm is used it can only give accurate results.



The categories, Of attacks	Attack exist in the training data and testing
DOS	back, land, neptune, pod, smurf, teardrop
Probe	ipsweep, nmap, portsweep, satan
R2L	ftp_write, guess_passwd, imap,multihop, phf, warezmaster teardrop
U2R	buffer_overflow, loadmodule, perl, rootkit

Fig 3. Types of attacks

Dos attack: The attackers made the computing devices too busy and it makes inaccessible to users. Dos attacks are classified into types they are flooding services and crashing services. It mainly attacks on the legitimate users such as banking, commerce and media companies or government and trade organizations.

Probe attack: Probe attack is happened by observing the silicon chip. So the attackers can easily access the sensitive information of the users from the computing wires and networks.

R2L: In this attack attacker sends a set of packets to the server through the internet. The user may loss some sensitive information and the attacker gain access.

U2R: In this attack the attackers trying to gain information like a normal user. It main goal is to obtain a root access to a system.

For this project some machine learning techniques for dataset they are explained in the following chart.

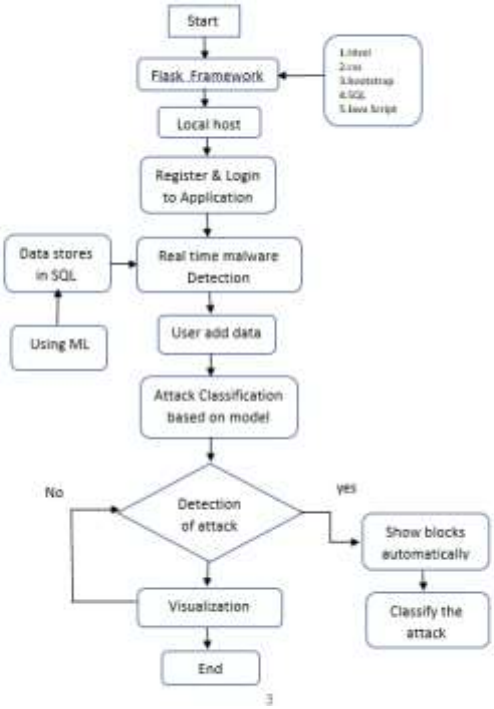


Fig 3. Architecture

Below diagram figure 4 explains the machine learning flow of the diagram.

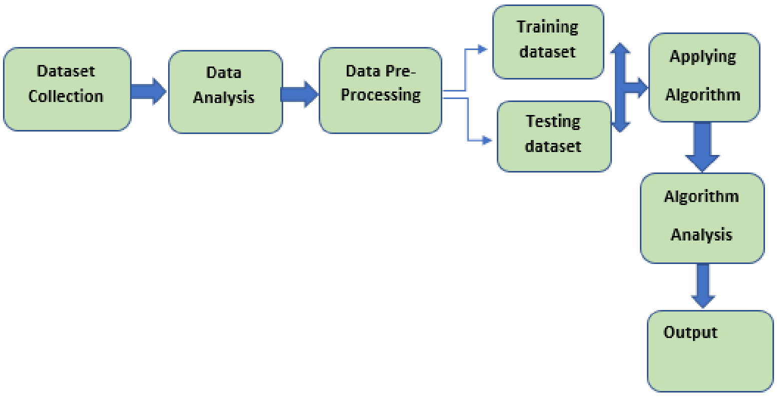
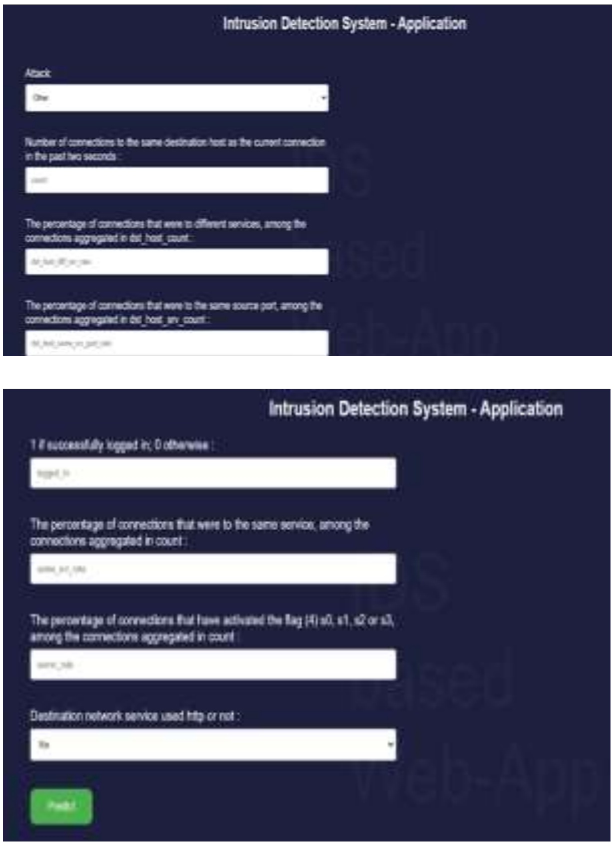


Fig 4 . Dataset

RESULT:



Intrusion Detection System - Application

Attack:

Other

Number of connections to the same destination host as the current connection in the past two seconds :

count

The percentage of connections that were to different services, among the connections aggregated in dst_host_count:

dst_host_diff_srv_rate

The percentage of connections that were to the same source port, among the connections aggregated in dst_host_srv_count :

dst_host_same_src_port_rate

Intrusion Detection System - Application

1 if successfully logged in; 0 otherwise :

logged_in

The percentage of connections that were to the same service, among the connections aggregated in count :

same_srv_rate

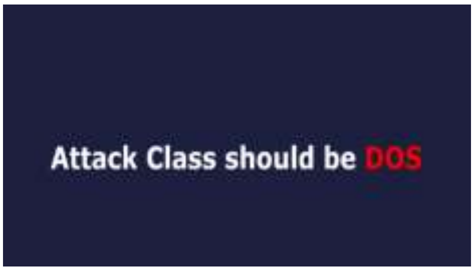
The percentage of connections that have activated the flag (4) s0, s1, s2 or s3, among the connections aggregated in count :

serror_rate

Destination network service used http or not :

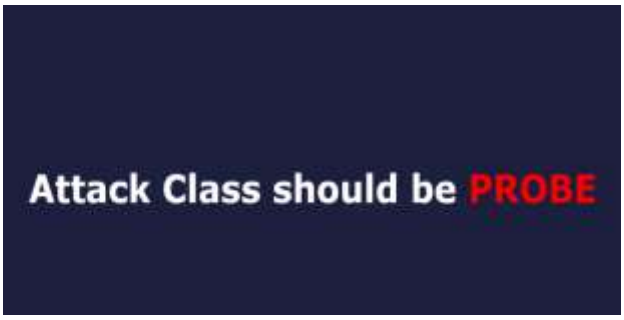
[illegible]

Fig 5 IDS Application



Attack Class should be DOS

Fig 6 Output1



Attack Class should be PROBE

Fig 7 Output 2

CONCLUSION:

We can easily predict the attack by using the machine learning techniques. Various algorithms such as support vector machine, Artificial neural network, Random forest, Decision tree are used to detect the attacks very effectively. We have implemented our proposed methodology and performed extensive evaluations. It will help to cyber security analysts to detect which attack is going on our network. Later they will use different techniques to recover from the loss. There is a need for customized models for security purposes.

FUTURE SCOPE:

In future we have to increase the efficiency of the method. Not only machine learning other technologies can also have merged to increase its accuracy. At the same time detection and removing attacks takes place once at a time projects will come into the picture.

REFERENCES:

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, “Port scanning techniques and the defense against them,” SANS Institute, 2001.

[3] M. Baykara, R. Das¸, and I. Karado ˘gan, “Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac¸larin incelenmesi,” in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] S. Staniford, J. A. Hoagland, and J. M. McAlerney, “Practical automated detection of stealthy portscans,” Journal of Computer Security, vol. 10, no. 1-2, pp. 105–136, 2002.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, “Surveillance detection in high bandwidth environments,” in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

[6] K. Ibrahimi and M. Ouaddane, “Management of intrusion detection systems based-kdd99: Analysis with lda and pca,” in Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017, pp. 1–6.

[7] N. Moustafa and J. Slay, “The significant features of the unsw-nb15 and the kdd99 data sets for network intrusion detection systems,” in Building Analysis Datasets and Gathering

Experience Returns for Security (BADGERS), 2015 4th International Workshop on. IEEE, 2015, pp. 25–31.

[8] L. Sun, T. Anthony, H. Z. Xia, J. Chen, X. Huang, and Y. Zhang, “Detection and classification of malicious patterns in network traffic using benford’s law,” in Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2017. IEEE, 2017, pp. 864–872.

[9] S. M. Almansob and S. S. Lomte, “Addressing challenges for intrusion detection system using naive bayes and pca algorithm,” in Convergence in Technology (I2CT), 2017 2nd International Conference for. IEEE, 2017, pp. 565–568.