Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 03, March -2018

# A SURVEY OVER DIFFERENT IDS SOFTWARE DEFINED NETWORKING APPROACH

Subhendu Kumar Gautam[1], Vivek Kumar[2], Imran Khan[3]

[1]*CSE Department, OIST College*
[2]*CSE Department, OIST College*

**Abstract —** *Data mining is an emerging technique today where there are number of responsible data storage and security aspects has been already applied , still there are some major issues which need to be discussed and explore based on the requirement in today's network and system, various frauds and intrusion related problems met and various requirement for such prevention system also been explored, here in this Synopsis we present an approach for intrusion detection and prevention system in system where the requirement is to deal with various anomaly and fraud detection system invention, here we propose a technique which explore and discuss about the intrusion detection system in cloud computing. The paper uses A DNN (Deep neural network) model for the anomaly and intrusion detection over the available dataset which is KDD (knowledge discovery dataset). A further extension to the work can be opting out in giving efficiency with deep learning and detection model. Also a further study is going to work on software defined networking with the more enhance and rich development with real-time dataset.*

***Keywords***- *IDS; Software outliers; DNN; KDD; Spamming detection; NID; Classification.*

## I. INTRODUCTION

Classification, regression and clustering are three approaches of data mining in which instances are grouped into identified classes. Classification is a popular task in data mining especially in knowledge discovery and future plan. It provides the intelligent decision making. Classification not only studies and examines the existing sample data but also predicts the future behavior of that sample data. It maps the data into the predefined class and groups. It is used to predict group membership for data instances. In Classification, the problem includes two phases first is the learning process phase in which for analysis of training data, the rule and pattern are created. The second phase tests the data and archives the accuracy of classification patterns.

Today in Dataset there exist data objects that do not comply with the general behavior or model of the data. Such data objects, which are heavy different from or inconsistent with the remaining set of data, are called outliers. An outlier is a data set which is different from the remaining data. Outlier is also referred to as deformity, deviants or anomalies in the data mining and statistics literature. In most applications the data is created by one or more generating processes, which could either reflect activity in the system or observations collected about entities. When the developing process behaves in a casual way, it results in the creation of outliers. Therefore, an outlier often contains useful information about anomaly characteristics of the systems and entities, which impact the data generation process. The recognition of such unusual characteristics provides useful application specific insights.

IDS have two approaches to analyze the traffic for intrusion: Misuse approach and Anomaly approach. Misuse detection based IDS follow well defined patterns (or signature) that can usually be detected by doing pattern matching on collected patterns from training data. This limits the number of false positives, but usually does not completely eliminate them all. Like virus scanners, Misuse based IDS cannot detect something that the network manager doesn't know about (i.e. new attack). For a Misuse based IDS to be useful, its signature sets must be constantly updated. Anomaly detection based IDS are based on observations of deviations from normal system usage patterns. They are detected by building up a normal profile of the system being monitored and detecting significant deviations from this profile. If Anomaly based IDS is set up with a narrow definition of normal, the IDS will generate large number of false positives.

Unlike Misuse based IDS, it can detect new attacks also. IDSs can be classified on basis of data source as Network based or Host based. The Network based IDS (NIDS) collect raw network packets as the data source from the network and analyze for signs of intrusions. Some examples of NIDS are NFR, Dragon, ISS, Snort, Bro, Cisco Secure IDS. Host based IDS (HIDS) operates on information collected from within an individual computer system such as operating system audit trails. C2 audit logs and System logs [4].

## II. RELATED WORK

Firewall and IDS both are related to network security. IDS differ from a firewall in that a firewall is primarily utilized as a traffic-filtering device, whereas An IDS is primarily utilized as a traffic–auditing device. An IDS looks for certain traffic patterns activity that has been determined as anomalous or possibly malicious. A firewall is designed to limit the access between networks in order to prevent intrusion. Generally, a firewall is not designed to provide detailed notification of attacks or anomalous network activity. An IDS evaluates a suspected event as it takes place or after it has

 

taken place, creates a detailed audit in one or more secured locations, signals an alarm or notifications. An IDS can also be setup to watch for attacks or events that originate from within an organization's network. This protects the organization from the possible legal punitive actions initiated as a result of attacks from inside the organization [17].

In Paper [17] an efficient SDN over the network is deployed and discussed, which make understanding of network flow. Network factors and cost analysis is also performed in this network.
An IDS cannot replace firewall and vice-versa because firewall does not have the intrusion detection capabilities of an IDS, an IDS does not generally have the firewalling capabilities of a firewall. The two technologies are very complementary and are generally deployed in union to achieve maximum level of security [16] , a classification and different level of IDS is depicted in figure 1 below.

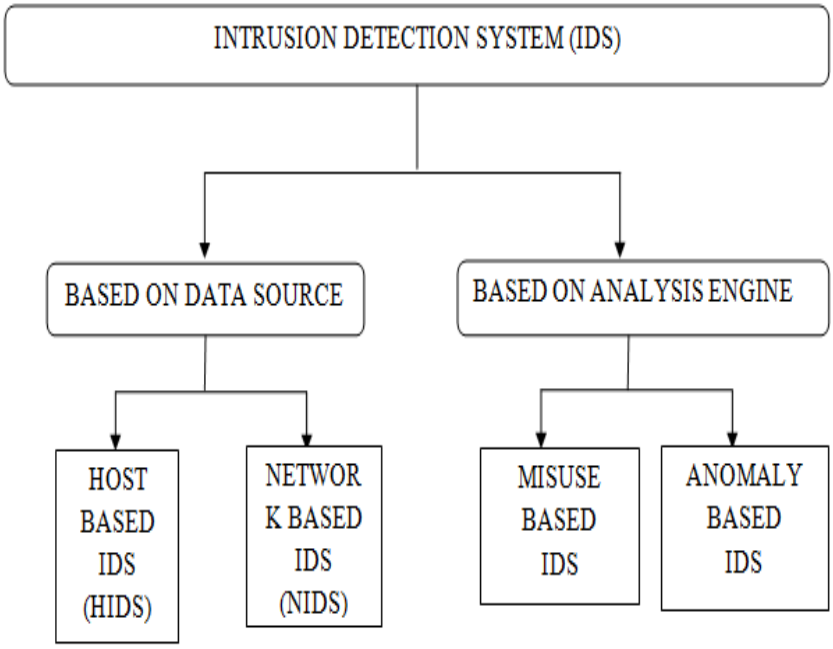


***Figure 1: IDS taxonomy for usage and availability.***

Having IDS to complement a firewall can provide an extra layer of protection to a system such as

- Identifying attacks that a firewall legitimately allows through (such as http attacks against web servers).
- Identify attempts such as port scan.
- Provides additional checks for holes/ports opened through firewall intentionally or unintentionally.

**DISADVANTAGES**

The current technique with the IDs are either host based or associated with the firewall , they have a low detection rate and accuracy. Also in terms of their parameters such as Recall and precision are not matching with the standard required in order to detection attack as an outlier and perform into the real-time entity.

**APPLICATION**

There are systems which are already using IDS system and still there are areas where IDS can be used which are mentioned below:

1. Military Network Area.
2. Traffic detection in a Network.
3. Firewall and Anti-virus system.
4. Fuzzy Search in Applications.
5. Image classification (Scene Detection).
6. Ranking classification.

## III. LITERATURE REVIEW

In this paper [1] propose decision tree based support vector machine which combines support vector machine and decision tree. It is an effective way for solving classification problems. This method decreases the training and testing time, increasing efficiency of the system. They include different ways to construct the binary trees by dividing the data

 

set into two subsets from root to the leaf until every subset consist only one class. Euclidean distance is used for measuring the separability between the classes.

In paper [2] examine the performance of framework for solving classification problem with support vector machine (SVM). The proposed paradigm construct binary tree for multi-class SVM, using criteria of natural classification: Homogeneity and Separation, with the aim of obtaining optimal tree [3]. This approach is more accurate in construction of tree. In the test phase, due to log complexity it is much faster than other methods that have problem of big class number. Here recognition rate was achieved by 57% on vowels of TIMIT corpus and 97.73% on MNIST dataset for 10 digits. Training time and number of support vectors are also reduced compared to other methods.
In paper [4] proposed a host based IDS using fuzzy inference rule that can detect changes in the hardware profile [4]. He used system performance log to evaluate accuracy and detection of system. With the help of deviation method, he selected effective features and generated fuzzy IF-THEN rules.

In Paper [5] first proposed in 1980 that security audit trails can play an important role in the security program for a computer system [6]. The purpose of his research was to improve the computer security auditing and surveillance capability of the customer's systems. Anderson postulated that it was possible to distinguish between a masquerader and a legitimate user by identifying deviation from historically-tracked system usage [7]. Denning proposed a model that was regarded as rule-based pattern matching system for audit records [8]. The model includes profiles for representing the normal behavior with respect to objects in terms of rules and any significant deviation from normal behavior termed as abnormal pattern. The generated audit records is matched with defined rules and checked for abnormal behavior.

In this paper [9] proposed a misuse detection system based on genetic algorithm approach [9]. She used Principal component analysis (PCA) to extract most important features. After that she used genetic algorithm to generate quality rules with the highest fitness values in every generation. Generated rules were used for classification of the intrusions and the normal connections in the testing data.

In this paper [10] proposed fuzzy clustering-Artificial neural network (FC-ANN) to enhance the precision and accuracy rate [10]. He divided the training data into different subset using fuzzy clustering. For each subset of training data, ANN was applied to learn the system precisely. After learning of each subset, he used fuzzy aggregation module to learn again and combine the different ANN's results. He demonstrated that proposed model gives better performance compared to BPNN, naïve Bayes.

In this paper [11] Madame inference system has been used to identify the accurate behavior of generated system log. Baghdad used five types of neural networks (NN) to determine which NN classifies well the attacks and leads to the higher detection rate of each attack [11]. He evaluated that among five NN that is multilayer perceptron (MLP), generalized forward (GFF), radial basis function (RBF), self-organization feature map (sofm), principal component analysis (PCA), GFF NN leads to the best confusion matrix in the multiclass case (DOS, U2R, R2L, and Probe). For same case, RBF performs the higher detection rate of DOS attack class. In case of single class (Normal or attack), PCA NN performs the higher detection rate.

In paper [12] discuss an intelligent alert clustering model for network based intrusion detection system. They proposed a novel integration of improved unit range (IUR), Principal component (PCA) and expectation maximization (EM) to detect intrusion alerts and filter out the unwanted alerts. To filter out unwanted or false positive alerts, they assigned a level of severity (high risk, medium risk, low risk) to each alerts. Hlaing used fuzzy decision tree classifier to detect intrusion alerts in network traffic [13]. She selected 10 best features by using mutual correlation feature selection algorithm and then applied fuzzy C4.5 decision tree algorithm on the training dataset. Experimental result showed that her proposed system achieved 99% classification accuracy. In [14],
In paper [15] proposed a way to use genetic algorithm to improve support vector machines (SVM) based Intrusion detection system (IDS). Integration of GA and SVM selects the optimal parameters as well as optimal feature set for dataset. They used KDD'99 dataset for evaluating results.
A complete study about the different technique over the mining and IDS approach shows the different algorithm performed work over intrusion detection over software defined networking dataset.

## IV. CONCLUSION

Data classification and IDS detection over large data is further be required research today. Study have been surveyed the techniques which are been performed for the outlier detection and in order to work with the Intrusion detection system with the data set provided. We have taken the various paper of different latest author and understood the work performed by them and finally we have concluded the discussion related to the previously performed algorithm in IDS system. Upon discussion of previous algorithm we are able to understand the further work which can be performed for outlier detection in KDD dataset.

 

**REFERENCES**

[1] M. S. Hoque, M. A. Mukit, M. A. N. Bikas; "An Implementation of Intrusion Detection System Using Genetic Algorithm"; IJNSA; vol. 4; 2012.
[2] Christopher M. K., Curtis E. Dalton, T. E. Osmanoglu, "Security Architecture: Design, Deployment and operations," RSA PRESS, Tata McGraw-Hill Edition: 2003.
[3] Vera MarinovaBoncheva, "A short survey of Intrusion Detection System," Bulgarian academy of sciences, Problems of Engineering Cybernetics and Robotics, Vol. 58; 2007, PP. 23-30.
[4] Om H., Gupta A. K., "Design of Host based Intrusion Detection System using Fuzzy Inference Rule," International Journal of Computer Applications, vol. 64(9), PP. 39-46, 2013.
[5] Stallings W., "Cryptography and Network Security Principles and Practices," Prentice Hall, 1998.
[6] J.P. Anderson, "Computer Security Threat Monitoring and Surveillance Technical report," James P. Anderson Co., Fort Washington, PA, April 1980.
[7] D. E. Denning, P. G. Neumann, "Requirement and model for IDES – A real time intrusion detection system," Computer Science Laboratory, SRI International, Menlo park, Technical Report # 83F83-01-00, 1985.
[8] D. E. Denning, "An Intrusion Detection Model," IEEE Trans. on Software Engineering, Vol.-SE-13(2), pp. 222-232, 1987.
[9] Bankovic Z., Stepanovic D., Bojanic S., Octavio N., "Improving network security using genetic algorithm approach," Elsevier, Computers and Electrical Engineering, Vol. 33, pp 438-451, 2007.
[10] Wang G., Hao J., Ma J., Huang L., "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering," Elsevier, Expert system with Applications Vol. 37, pp 6225-6232, 2010.
[11] Beghdad R., "Critical study of neural network in detecting intrusions," Elsevier, Computers and Security, Vol. 27, pp. 168-175, 2008.
[12] Siraj M. M., Maarof M. A., Hashim S.Z.M., "Intelligent Alert Clustering Model for Network Intrusion Analysis," International J. Advance Soft Comput. Appl. Vol. 1(1), pp. 33-48, 2009.
[13] Hlaing T., "Feature Selection and Fuzzy Decision Tree for network Intrusion Detection," International journal of Informatics and Communicational Technology, Vol. 1(2), pp. 109-118, 2012.
[14] Kim D. S., Nguyen H., Park J. S., "Genetic Algorithm to Improve SVM Based Network Intrusion Detection System," Proceedings of the 19th International Conference on Advance Information Networking and Applications, Vol. 2, pp. 155-158, 2005.
[15] http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/data/1998data.htm
[16] Ali Malik; Benjamin Aziz; Mo Adda; Chih-Heng Ke, " Optimisation Methods For Fast Restoration of Software-Defined Networks", 2017, IEEE.
[17] RaphaelHorvath, DietmarNedbal, MarkStieninger, "A Literature Review on Challenges and Effects of Software Defined Networking", Elsevier 2016.

 