Impact Factor (SJIF): 3.632

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 3, Issue 3, March-2016

Intrusion Detection System

Harendra Vaidya [1], Shahrukh Mirza [2], Nayan Mali [3]

Student, IT, Sigma Institute of Engineering, Vadodara [1][2]
Assistant professor, IT, Sigma Institute of Engineering, Vadodara [3]

Abstract *— Intrusion detection system is process of detecting intrusion in database, network or any other device for providing secure data transmission. With recent advance in network based technology and increased dependability of our everyday life. Privacy is major part in networking and detection of intrusion based security attacks play major role in network. In network selection system is very important for increase analysis effort and process time. Data mining is effective feature of selection and intrusion detection in network.*

Keywords- *Intrusion, Ad-hoc network, Network Security, Classification, Intrusion Detection System.*

1. INTRODUCTION

Now a day, Communication technology has been widely used for the communication and the transmission purpose in many applications and wide usage for this security became a challenging problem for this current technology. Intrusion Detection Systems (IDSs) are an adjective component of network security. They collect and check packets, looking for unnecessary packet and their behaviors. As soon as unnecessary event is detected, security process takes actions for deleting or repair the system. Main focus is on detecting known and unknown attacks in fast networks in order to reduce attack by shrinking the time gap between the real attack and its detection. This paper contribution is to build two levels IDS in order to detect abnormal behavior of network traffic and covering fast networks

1.1 Data Mining

The process of extracting useful and previously unnoticed model or pattern are large data set. Data mining is a process to find useful information form data. Data mining can be used to solve security problem as below reasons.

- To process huge amount of data.
- Analysis, extracting and compressing specific data.

Intrusion: Any unauthorized person not permitted attempt to access damage or malicious use of via information resources.

Intrusion Detection: Detection of steal and break-in attempts through automated software systems.

Intrusion Detection System: Security systems, which detect and possibly stop intrusion detection activities. Intrusion is some time also called as hacker or cracker attempting to destroy or misuse your system. in 1980, beginning concept of intrusion detection, defined as software application which can monitors networks and system activities for detect unauthorized users. IDS tools aim to detect computer attacks, all inbound and outbound network activity, recognize trusting patterns that may indicate a network or system, monitor attacker who want crash systems(hacker). An IDS system is like alerting alarm system in house it alert malicious activates.

1.2 Types of IDS

---NIDS (Network Intrusion Detection Systems)
---HIDS (Host Intrusion Detection Systems)

1.2.1 Network Intrusion Detection Systems: A network IDS (NIDS) monitoring or capture all traffic. this place on the network are divide. Network Intrusion Detection involve looking at the packets on the network as they pass by the NIDS. Packets are considered to be of interest if they match a signature or certain behavior [2][6][7]. Network Intrusion Detection Systems are kept at a important point or points within the inbound and outbound network to monitor traffic to and from all devices.

1.2.2 Host Based Intrusion Detection Systems: Host Based Intrusion Detection involves not only looking at the network traffic and out of single computer, but also checking the all of your system files and watching some unwanted processes. To get complete coverage at your network with HIDS, you must load the software on every computer [1][6]. Host based Intrusion Detection is more and more effective in detecting insider attacks than is NIDS. HID Systems are run

on different hosts or devices on the network. A HIDS monitoring the inland and outland packets from the device only and will inform the user or administrator of malicious activity is detected.

2. EXISTING SYSTEM

2.1 KDDCUP'99 Data Set

The data set used to perform the testing is taken form KDD Cup'99 which is widely accepted as a benchmark dataset and identify by many researchers. 10% of KDD cup '99 from KDD Cup'99 data set was chosen to evaluate rules and testing data sets to detect intrusion. The all KDD Cup'99 data set contains 41 features. Connection are labels as normal or attacks fall into four main categories [1][2].

1. DOS -Denial of service
2. Probe -port scanning
3. U2R -unauthorized access to root person
4. R2L -unauthorized remote login to machine

In this dataset there are 3 group of features: Basic, content based, time based features.

Training set consists 5 million connections.

- 10% training set -494,021 connections
- Test set have -311,029 connections

Test data has attacks types that are not present in training data. Problem is more original:

-Prepared set contains 22 attack types.

-Test data contains additional 17 new attack types that belong to one of four main categories.

The ``same host" features examine only the connections in the extinct two seconds that have the same destination to the host as the current connection, and calculate statistics related to protocol work, service, etc.

The same service features examine only the connections in the pass two seconds that have the similar service as the present connection. "Similar host" and "similar service" features are is called time based traffic features of the connection records.

Some effected attacks scan the hosts using a much larger time delay than two seconds, for example once per minute. Therefore, connection records were also sorted by receiver host, and features were constructed using a window of 100 connections to the same host instead of a time window. This result a set of so called host based traffic features.

Unlike most of the DOS and probing attacks, there appears to be no sequence patterns that are random in records of R2L and U2R attacks. This is because the DOS and probing attacks interconnect many connections to some host in a very small period of time, but the R2L and U2R attacks are embedded in the data part of packets, and normally involve only a single connection.

Useful algorithms for mining the un-design data portions of packets automatically are an open research question. used domain knowledge to add features that look for doubtful behavior in the data portions, such as the number of failed login attempts. These features are called related features.

A complete listing of the set of features express for the connection records is given in the tables below. The data schema of the contest dataset is available in machine-readable form.

The proposed algorithm for KDDCup'99 data classification in order to predefine our criteria. In this model first different preprocessing stage is applied after they apply our new access for segment.

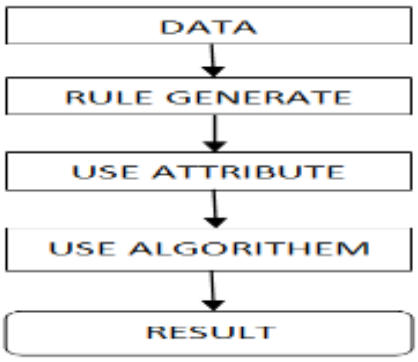

Fig.2.1 process steps

3. STUDY

Tp rate TPR=TP/(TP+FN)
Fp rate: FPR=FP/(FP+TN)
Tn rate (or specificity):
Precision = TP/(TP+FP)
Recall = TP/(TP+FN)
FNR=1-TPR
Accuracy = (TP + TN) / (TP + TN + FP + FN)
Correctly classified records = TP + TN
Incorrectly classified records = FP + FN

3.1 NAIVEBAYES:

```
Correctly Classified Instances       14729               89.3967 %
Incorrectly Classified Instances      1747               10.6033 %
Kappa statistic                          0.7866
Mean absolute error                      0.1052
Root mean squared error                  0.318
Relative absolute error                 21.129  %
Root relative squared error             63.7554 %
Total Number of Instances            16476

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.908     0.122      0.895     0.908     0.902      0.968    normal
                 0.878     0.092      0.892     0.878     0.885      0.964    anomaly
Weighted Avg.    0.894     0.108      0.894     0.894     0.894      0.966

=== Confusion Matrix ===

    a    b   <-- classified as
 8015  811 |    a = normal
  936 6714 |    b = anomaly
```

Fig.3.1 Naive Bayes Result

3.2 NB Updateable:

```
Time taken to build model: 1.09 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances       49108               89.4173 %
Incorrectly Classified Instances      5812               10.5827 %
Kappa statistic                          0.7871
Mean absolute error                      0.1052
Root mean squared error                  0.3186
Relative absolute error                 21.1185 %
Root relative squared error             63.8527 %
Total Number of Instances            54920

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.914     0.128      0.89      0.914     0.902      0.968    normal
                 0.872     0.086      0.899     0.872     0.885      0.964    anomaly
Weighted Avg.    0.894     0.108      0.894     0.894     0.894      0.966

=== Confusion Matrix ===

     a     b   <-- classified as
 26664  2517 |     a = normal
  3295 22444 |     b = anomaly
```

Fig. 3.2 NB Updateable Result

3.3 Result OF Two Algorithms

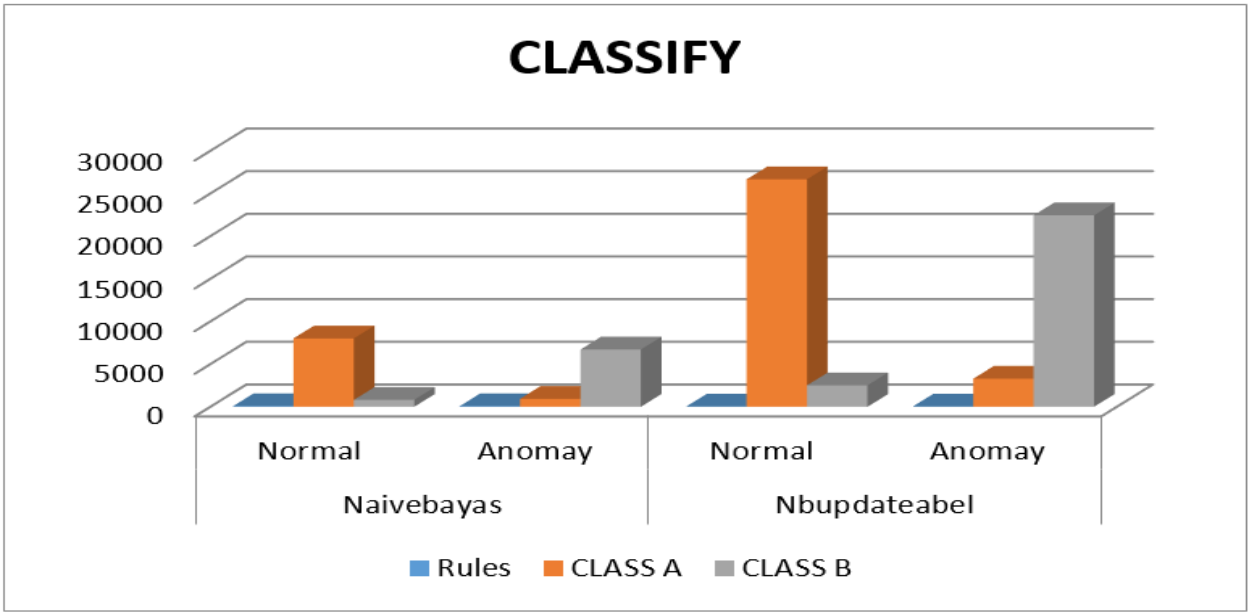

Fig.3.3 combination of two algorithm

4. CONCLUSION

Data mining is improve intrusion based security attacks detecting system by adding a new level of detection of network data indifferences. It is highly required to identify appropriate features to categorize into different types of attack. Feature selection abbreviate the size of network data which improve finally performance of intrusion detection system. Naive Bayes algorithm which is used for experimentation is efficient algorithm of feature selection in KDD cup 99 dataset. This feature identification helps to improve capacity of intrusion detection system.

5. REFERENCES

[1] N.S.Chandolikar, V.D.Nandavadekar, "Selection of Relevant feature for Intrusion Attack ClassificationBy Analyzing KDD Cup 99.MIT International journal of Computer Science & Information Technology, ISSN No.2230-7621 MIT Publication.

[2] Bakshi, A. & B. Yogesh. 2010. Securing Cloud from DDOS Attacks Using Intrusion Detection System in irtual Machine. In Second International Conference on Communication Software and Networks. pp. 260–264.

[3] C. C. Lo, C. C. Huang & J. Ku. 2008. Cooperative Intrusion Detection System Framework for Cloud Computing Networks. In First IEEE International Conference on Ubi-Media Computing. pp. 280–284.

[4] C. Mazzariello, R. Bifulco & R. Canonoco. 2010. Integrating a network IDS into an Open source Cloud computing. In Sixth International conference on Information Assurance and Security (IAS). pp. 265–270.

[5] Jeffrey L U, Anupam J. Next Generation Data Mining, MIT Press, 2003

[6] Ms. Vinita Shrivastava, Mr. Neetesh Gupta "Performance Improvement Of Web Usage Mining By Using Learning Based K-Mean Clustering" International Journal of Computer Science and its Applications

[7] S.Taherizadeh and N.Moghadam "Integrating web content mining into web usage mining for finding patterns and predicting user's behaviors", An International journal of information science and management, January / June- 2009, Vol.7.

[8] Prakash S Raghavendra, Shreya Roy Chowdhury, SrilekhaVedulaKameswari "Web Usage Mining using Statistical Classifiers and Fuzzy Artificial Neural Networks" International Journal Multimedia and Image Processing (IJMIP), Volume 1, Issue 1, March 2011.