Comparative analysis of different Categories of Anomaly Detection System.

Rshma Chawla[1], Gurpreet Kaur[2]

Assistant Professor, MMICT&BM, MMU, Mullana (Ambala*) India*

Abstract

Intrusion detection is so much popular since the two decades where intrusion attempted to break into or misuse the system. This IDS system has ability to detect the virus, malware, spy ware &different form of viruses.IDS classifying the two categories one is structure based &another is according to detection techniques. The network based &host based are included in structure based IDS. The detection techniques are used to detect suspicious activity both at anomaly based and signature based. The general scheme of anomaly preprocessor & its types of categories are explained in this paper.

Keyword:-*Anomaly, Anomaly detection techniques, common anomaly based IDS, Intrusion Detection System, Signature based IDS.*

1. Introduction

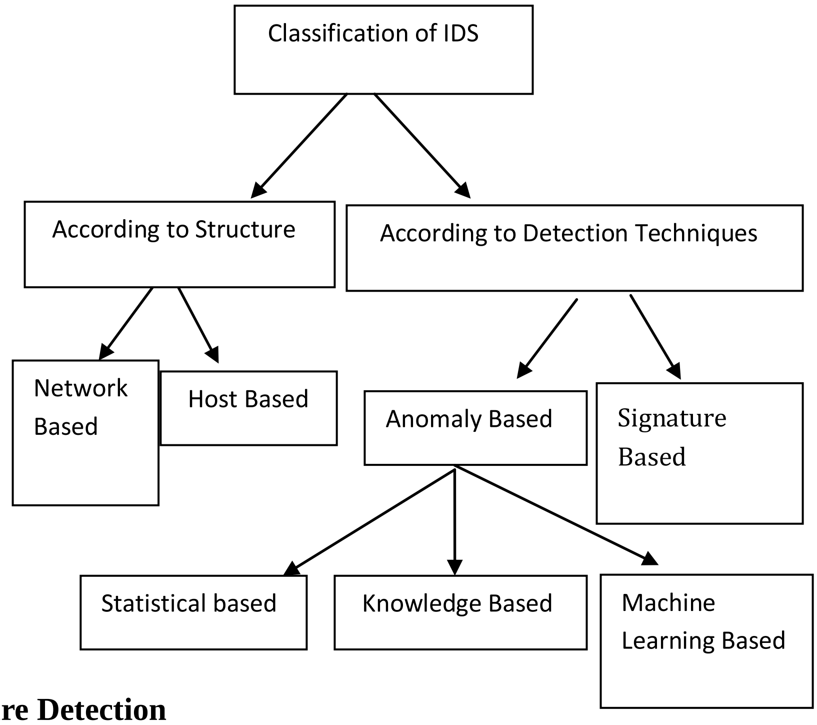
1.1Intrusion Detection System

Intrusion detection systems are security tools like other measures such as antivirus software, firewall etc. It proposed to improve computer security because it is not feasible to build completely secure systems [23]. In particular IDSs are used to identify, assess and report unauthorized network activities, so that appropriate actions may be taken to prevent any future damage. Intrusion detection systems are the 'burglar alarms' of the computer security field [3]. The aim is to defend a system by using a combination of an alarm that sounds whenever the site's security has been compromised and an entity –most often a site security officer (SSO) that can respond to the alarm and take the action. There is often the need to update an installed Intrusion Detection System (IDS) due to new attack methods or upgraded computing environments. Since many current IDSs are constructed by manual encoding of expert knowledge, changes to IDSs are expensive and slow. Intrusion detection systems are classified into anomaly based or signature based. Signature based uses specifically known patterns of unauthorized behavior to predict and detect subsequent similar attempts [23]. It is also known as misuse detection. These specific patterns are called signature. On the other hand, anomaly based detectors attempt to estimate the "normal "behavior of the system to be protected and generate an anomaly alarm. Another possibility is to model the "abnormal" behavior of the system and to raise an alarm when the difference between the observed behavior and the expected one falls below a given limit.

1.2Types of Intrusion Detection System

IDS can also be categorized according to the detection approaches they use. Basically, there are to detection methods: anomaly detection and signature detection [25]. The major difference between the two methods is that anomaly detection analyzes the properties of normal behavior while signature detection identifies intrusion based on features of known attacks. Signature based uses specifically known patterns of unauthorized behavior to predict and detect subsequent similar attempts [2]. It is also known as misuse detection. These specific patterns are called signature. On the other hand, anomaly based detectors attempt to estimate the "normal "behavior of the system to be protected and generate an anomaly alarm [20]. The following subsections explain the two detection approaches.

Fig1. Classification of IDS



1.2.1Signature Detection

Any action that conforms to the pattern of a known attack is measured in it. The main issues in signature detection system are now to write a signature that encompasses all possible variations of attack .In contrast do not match non-intrusion activity [21].It identifies intrusion by matching monitored events to patterns. These specific patterns are called signatures. For host-based intrusion detection one example of signature is "Three failed logins". For network intrusion detection, a signature can be as simple as a specific pattern that matches a portion of a network packet. The major advantage is that the high accuracy in detecting known attacks. However, its detection ability is limited by signature database [25]. Otherwise new attacks are transformed into signature and added to database. Signature IDS cannot detect any attack of this.Diffrent techniques such as expert system, signature analysis and state transition analysis are utilized in misuse detection.

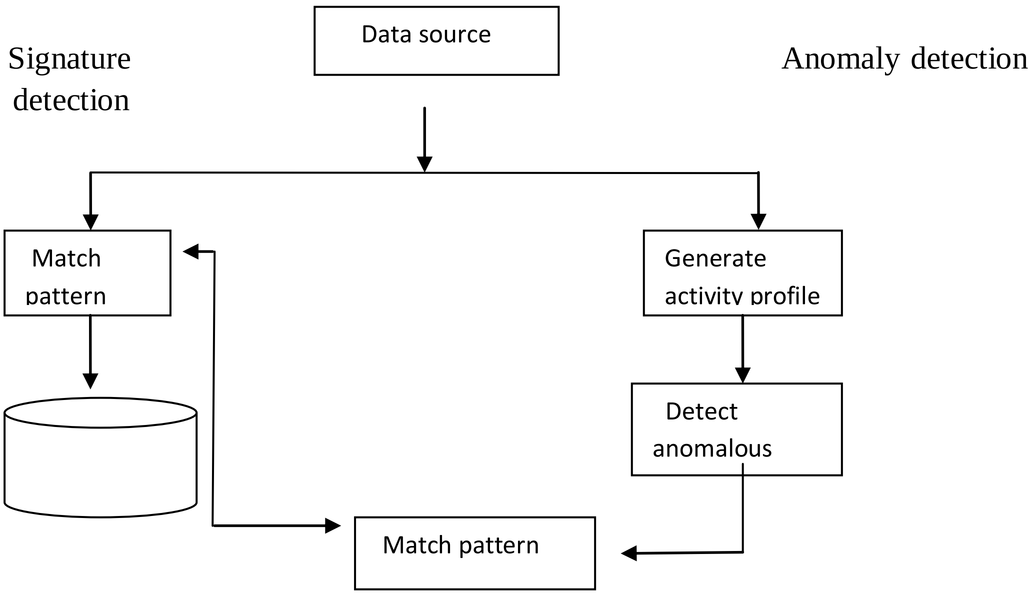


Fig2.General operation of IDS

1.2.1 Anomaly Detection

It is based on the normal behavior of the subject. Any action that significantly deviates from the normal behavior is considered as intrusive. An example: if a user logs on and off of machine more than 10 times a day, instead of the normal 1 or 2.Also if a computer is used at 2:00 am when normally no one outside of business hours should have access, this should raise some suspicious [8, 2].At another level anomaly detection can investigate user patterns such as profiling the program executed daily. If a user in the management department suddenly starts accessing programs or compiling code the system can properly alert its administrator [8].

2. Related Work.

Asmaa Shaker Ashore [2] examines the importance of Intrusion Detection Systems its categories and a classification .It also concludes that IDS is basically detects attacks signs and then alerts. In terms of performance, IDS becomes more accurate as it detects more attacks raises fewer positive alarms.

P.Garcia-Teodoro [23] discusses the main A-NIDS technologies, together with their general operational architecture and provides a classification for according to the type of processing related to the “behavioral model for the target system. The main features of several currently available IDS systems platforms.

J. Gomez [15**]** presents a new anomaly pre-processor that extends the functionality of Snort IDS, making it a hybrid IDS. It has been verified that when the number of elements increases it has less sensitivity and detect few attacks. Results also denote the importance of training the system during a long time to reduce the number of false alarms.

V. Jyothsna,"A [27**]** presents anomaly-based approaches are efficient, signature-based detection is preferred for mainstream implementation of intrusion detection systems. High detection rate of 98% at a low alarm rate of 1% can be achieved by using these techniques. It elaborates the main anomaly based network intrusion detection technologies along with their operational architectures and also presents a classification based on the type of processing that is related to the "behavioral" model for the target system**.**

Mueen Uddin [18] explains the IDS system is how to keep up with large volume of incoming traffic. When each packet needs to be compared with every signature in database. In contrast compare with anomaly detection technique. It introduced a new model of dynamic multilayer signature based IDS.

Augustine Soule [3**]** explains here how any anomaly detection method can be viewed as a problem in statistical hypothesis testing. It compares different method for analyzing residual. These methods focus on different aspects of the traffic pattern change.

Bar ford at al[6] presented a frame work for detecting and localizing performance anomalies based on using an active – probe-enabled measurement infrastructure deployed on periphery of a network. Their frame work has three components; an algorithm for detecting performed anomalies on a path, algorithm for selecting which paths to probe at a given time in order to detect performance anomalies and an algorithm for identifying the links that are causing an identified anomaly on path (i.e. localizing).

3. Introduction to Anomaly

3.1What is Anomaly

Anomalies can be treated as pattern not observed before. [23] Anomaly Detection refers to detection patterns in a given data set that do not conform to an established normal behavior. The patterns thus detected are called anomalies [16]. It translates to critical and actionable information in several application domains. Anomaly Intrusion detection system is ineffective in detecting insider attacks, an intrusion detection system that employs only one of that method will have a limited range of intrusions .The main benefit of anomaly-based detection techniques is their potential to detect previously unseen intrusion events.Fig3.Shows the general scheme of anomaly detection module .Using two different operation modes: training mode and anomaly detection mode. Using the training mode the system records in a database the network traffic considered as normal and expected. Both operation modes share the same functionality. When the pre-processor of Snort receive a package, it is classified according to its class. (if the package is primary/secondary and if the package belongs to a network server or a client) and it stores the vector-class package, i.e. the system is recording and counting the network traffic. When the system is in training mode it stores the recorded information in the database and later it obtains a profile of the normal activity. The information store in the database is used when the system is in detection mode. Daily and each time the system is executed the activity profiles of the most active clients and servers in the network are loaded from the database. Therefore as the expected traffic is recorded in the database and compared with the real traffic passing through the network. If it is detected a deviation in the traffic higher than a certain percentage it means that something abnormal is happening and an incidence of abnormality is registered by the system

3.2Anomaly Detection Categories

Anomaly detection is based on a host or network. Many different techniques are used based on type of processing related to behavioral model [23].They are: Statistical based, knowledge model, Machine Learning based. [16]. (See Fig. 4)

1) Statistical-based

The behavior of the system is represented from a random viewpoint. In statistical-based techniques, the network traffic activity is captured and a profile representing its behavior is created. This profile is based on metrics such as the traffic rate, the number of packets for each protocol, the rate of connections, the number of different IP addresses, etc. As the network events occur, the current profile is determined [21].

Operational Model (or) Threshold Metric

The count of events that occur over a period of time determines the alarm to be raised if fewer then m or more than n events occur. This can be visualized in Win2k lock, where a user after n unsuccessful login attempts here lower limit is ‘0’and upper limit is n Executable files size downloaded is restricted in some organizations about 4MB.The difficulty in this sub model is determining ‘m’and ‘n’.

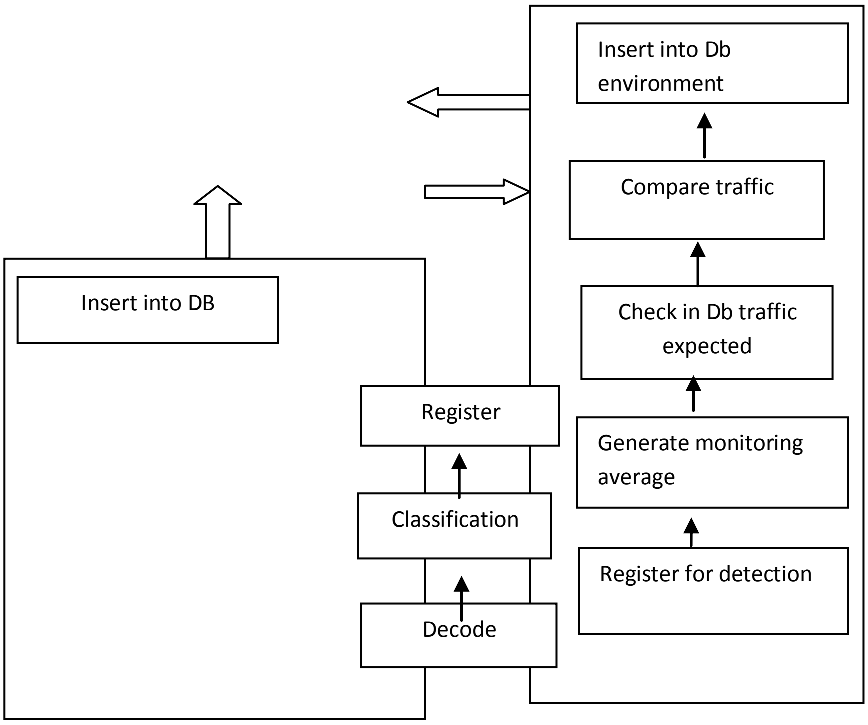

Insert into Db environment
Compare traffic
Check in Db traffic expected
Generate monitoring average
Register for detection
Insert into DB
Register
Classification
Decode

Fig3.General Scheme of Anomaly Pre-processor

Markov Process or Marker Model

The Intrusion detection in this model is done by investigating the system at fixed intervals and keeping track of its state; a probability for each state at a given time interval. The change of the state of the system occurs when an event happens and the behavior is detected as anomaly if the probability of occurrence of that state is low.

Statistical Moments

In statistical mean, standard deviation, or any other correlations are known as a moment. If the event that falls outside the set interval above or below the moment is said to be anomalous. The system is subjected to change by considering the aging data [13] .There is two major advantages over an operational model. First, prior knowledge is not required determining the normal activity in order to set limits; Second, determining the confidence intervals depends on observed. User data, as it varies from user to user. Threshold model lacks this flexibility. The major variation on the mean and standard deviation model is to give higher weights for the recent activities.

Multivariate Model

The major difference between the mean and standard deviation model is based on correlations among two or more metrics. If experimental data reveals better judicious power can be achieved from combinations of related measures rather than treating them individually [13].

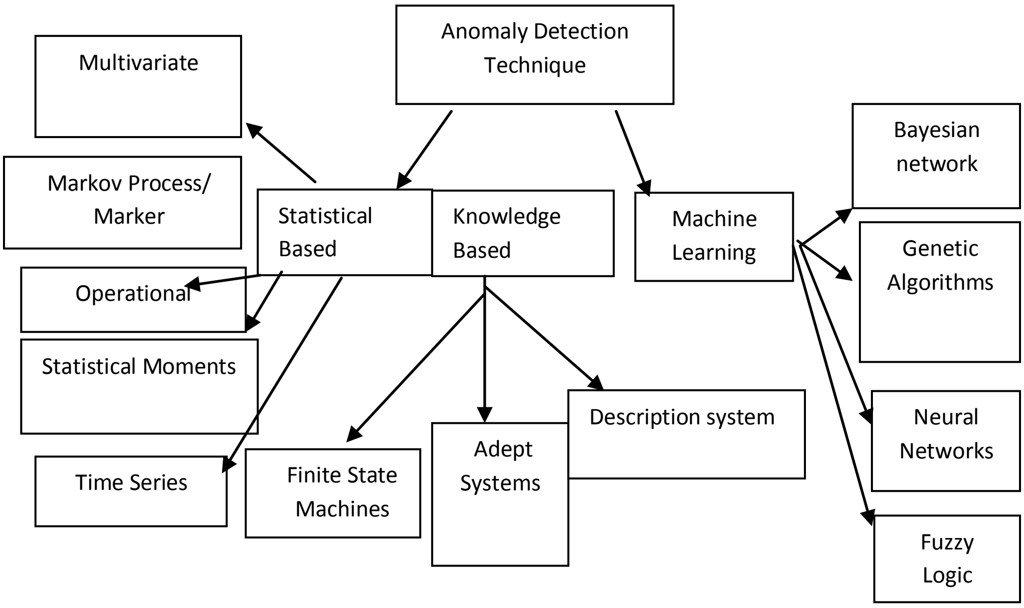


Fig4.Classification of Anomaly Based Intrusion Detection

Time Series Model

Interval timers together with an event counter or resource measure are major components in this model. Order and inter-arrival times of the observations as well as their values are stored [25]. If the probability of occurrence of a new observation is too low then it is considered as anomaly. The disadvantage of this model is that it is more computationally expensive.

2) Knowledge based

Expert system approach is one of the most widely used knowledge-based IDS schemes. Expert systems are intended to classify the audit data according to a set of rules [23], involving three steps. First, different attributes and classes are identified from the training data. Second, a set of classification rules, parameters or procedures are deduced. Third, the audit data are classified accordingly. More restrictive/particular in some senses are specification-based anomaly methods, for which the desired model is manually constructed by a human expert, in terms of a set of rules (the specifications) that seek to determine legitimate system behavior. If the specifications are complete enough [3], the model will be able to detect illegitimate behavioral patterns. Moreover, the number of false positives is reduced, mainly because this kind of system avoids the problem of harmless activities, not previously observed, being reported as intrusion.

Finite State Machine

A finite state machine (FSM) or finite automation is a model of behavior captured in states, transitions and actions [25]. A state contains information about the past, i.e. any changes in the input are noted and based on it transition happens [19]. An action is a description of an activity that is to be performed at a given moment. There are several action types: entry action, exit action, and transition action.

Adept Systems

Human expertise in problem solving is used in adept systems. It solves uncertainties where generally one or more human experts are consulted. These systems are efficient in certain problem domain, and also considered as a class of artificial intelligence (AI) problems [13]. Adept Systems are trained based on extensive knowledge of patterns associated with known attacks provided by human experts.

Description Scripts

Numerous proposals for scripting languages, which can describe signatures of attacks on computers and networks, are given by the Intrusion Detection community [13]. All of these scripting languages are capable of identifying the sequences of specific events that are indicative of attacks.

3) Machine Learning Based Detection Techniques

Machine learning techniques are based on establishing an explicit or implicit model that enables the patterns analyzed to be categorized. A singular characteristic of these schemes is the need for labeled data to train the behavioral model, a procedure that places severe demands on resources. The main advantages are Flexibility and adaptability, Capture of interdependencies [21]. Their main drawback is High dependency on the assumption about the behavior accepted for the system. There has some sub types Markov models (stochastic Markov theory), Neural networks (human brain foundations), Fuzzy logic (approximation and uncertainty), Genetic algorithms (evolutionary biology inspired), Clustering and outlier detection (data grouping).

Bayesian networks

A Bayesian network is a model that encodes probilistic relationships among variables of interest. This technique is generally used for intrusion detection in combination with statistical schemes, a procedure that has several advantages[15], including the capability of encoding interdependencies between variables and of predicting events, as well as the ability to incorporate both prior knowledge and data.

Genetic algorithms

They are a particular class of evolutionary algorithms that use techniques inspired by mutation, selection and recombination. These algorithms constitute another type of machine learning –based technique, capable of deriving classification rules and / or selecting features parameter for the detection process [16].The main advantage of this algorithm is the use of a flexible and robust global search method that coverage to a solution from multiple directions. The disadvantage is the resource consumption is involved in it.

Neural

Neural network have been adopted in the field of anomaly intrusion detection, mainly because of their flexibility and adaptability to environment changes. This detection approach has been employed to create user profiles [13]to identify the intrusive behavior of traffic patterns[8] etc. common characteristics is that they do not provide a descriptive model that why a particular detection a decision has been taken.

Fuzzy logic techniques

Fuzzy logic is derived from fuzzy. Fuzzy techniques are thus used in the field of anomaly detection mainly because the features to be considered can be seen as fuzzy variables.[7] This kind of processing schemes considers an observation as normal if it lie with in a given interval.[12]

Clustering

It distances measure. The procedure most commonly used for this consists in selecting are representives point for each cluster. Then, each new data point is classified as belonging to a given cluster according to the proximity to the corresponding representative point [16]. Some points may not belong to any cluster; these are representing the anomalies in the detection process [5].

4. Conclusion

The present paper discusses the foundations of main intrusion detection system technologies. All anomaly – based intrusion detection systems works on the assumption that normal activities differ from the abnormal activities (intrusions) substantially. It provides classification of intrusion detection system. The categories of anomaly technique are elaborate in it.

References

[1]Asmaa Shaker Ashoor," *Importance of Intrusion Detection System (IDS)",* International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.4, October 2010

[2]Asmaa Shaker Ashoor, "*Importance of Intrusion Detection System (IDS*)", International Journal of Scientific & Engineering Research, Volume 2, Issue 1, ISSN 2229-5518 January-2011

[3]Augustin Soule**,** "*combining Filtering &statistical methods for anomaly detection"*, USENIX association, internet measurement conference 2005.

[4]Barbara, D., Couto, J., Jajodia, S., & Wu, N., *"an Architecture for anomaly detection"*. In D.Barbara & S. Jajodia (Eds.), Applications of Data Mining in Computer Security (pp. 63--76). Boston: Kluwer Academic. (2002).

[5] Barnett V, Lewis T." *Outliers in statistical data*", Wiley, ISBN 9780471930945; 1994.

[6] Bridges S.M., Vaughn R.B. Fuzzy data mining and genetic algorithms applied to intrusion detection. In: Proceedings of the National Information Systems Security Conference;.p.13–31,2000

[7]Cansian A.M., Moreira E**.,** Carvalho A., Bonifacio J.M. "*Network intrusion detection using neural networks"*. In: International Conference on Computational Intelligence and Multimedia Applications (ICCMA'97); 1997. p. 276–80.

[8]Denning ED." *An intrusion-detection model"*. IEEE Transactions on Software Engineering 1987; 3(2):222–32.

[9] Debar H., Becker M., Siboni, D.*" A neural network component for an intrusion detection system".* In: IEEE Symposium on Research in Computer Security and Privacy;. p. 240–50. 1992

[10]Dickerson J.E. "*Fuzzy network profiling for intrusion detection*". In: Proceedings of the 19th International Conference of the North American Fuzzy Information Processing Society (NAFIPS); p. 301–6., 2000.

[11]Fox K., Henning R., Reed J., Simonian, R. "*A neural network approach towards intrusion detection*". In: 13th National Computer Security Conference; 1990. p. 125–34.

[12]Heckerman D. "*A tutorial on learning with Bayesian networks",* Microsoft Research; 1995. Technical Report MSRTR-95-06.

[13] J. Gómez1 C. Gil2, N. Padilla1, R. Baños2, and C. Jiménez11Dpt "*Design of a Snort-Based Hybrid Intrusion Detection System"*, Lenguajes y Computation, Universidad de Almeria, Spain (pp. 515–522), 2009.

[14]K.Hanumantha Rao,"*Implementation of Anomaly Detection Technique Using Machine Learning Algorithm"*, International Journal of Computer Science and Telecommunications [volume 2, Issue 3, June 2011].

[15]Li Tian, *"Research on Network Intrusion Detection System Based on Improved K-means Clustering Algorithm"*, Computer Science-Technology and Applications, 2009. IFCSTA '09. International Forum

[16]Liu, Y., Tian, D., Wang, A., ANNIDS: "*Intrusion Detection System Based on Artificial Neural Network", P*roceedings of the 2nd International Conference on Machine Learning and Cybernetics, 2003

[17]Lazarevic A, Kumar V, Srivastava J." *Intrusion detection: a survey, managing cyber threats: issues, approaches, and challenges*". Springer Verlag; 2005. p. 330.

[18]Mueen Uddin, "*Dynamic Multi-Layer Signature Based Intrusion Detection System Using Mobile Agents"*, International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.4, October 2010

[19]McHugh J, the 1998 Lincoln laboratory IDS evaluation. A critique. In: RAID. LNCS, vol. 1907; 2000. p. 145–61.

[20]P.Garcia-Teodoro**,** "*Anomaly –Based Network Intrusion Detection: Techniques, Systems and challenge*", computer security 28, 18-28,(2009)

[21]Paul Barford**,** University of Wisconsin, Nick Duffield AT&T, Amos Ron University and Joel Sommers Colgate, "*Network Performance Anomaly Detection and Localization*" Info COM 2009.

[22]Rashmi Chaudhary,"*Survey of Network Intrusion Detection Using K-Mean Algorithm***"**ISSN: 2277 128X, Volume 3, Issue 3, March 2013

[23]Sequeira K., Zaki M. ADMIT," *anomaly-based data mining for intrusions*". In: Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2002. p. 386–95

[24]Staniford-Chen S., Tung B., Porrar P., Kahn C., Schnackenberg D., Feiertag R., et al. "*The common intrusion detection framework data formats*." Internet draft 'draft-staniford-cidf-dataformats-00.txt', 1998

[25]V. Jyothsna,"*A Review of Anomaly based Intrusion Detection Systems"*, International Journal of Computer Applications (0975 – 8887) Volume 28– No.7, August 2011.