Comparative Analysis of Anomaly Detection Approaches for IDS

Ruby Dahiya*

Abstract

With the explosive growth of the Internet and the increased availability of tools for attacking networks, intrusion detection becomes a critical component of network administration. Intrusion detection systems gather information from a computer or network of computers and attempt to detect intruders or system abuse [16].Generally, an intrusion detection system will notify a human analyst of a possible intrusion and take no further action, but some newer systems take active steps to stop an intruder at the time of detection. There are two major intrusion detection techniques: misuse detection and anomaly detection. Misuse detection discovers attacks based on the patterns extracted from known intrusions. Anomaly detection identifies attacks based on the deviations from the established profiles of normal activities. Activities that exceed thresholds of the deviations are detected as attacks. Misuse detection has low false positive rate, but cannot detect new types of attacks. Anomaly detection can detect unknown attacks, under a basic assumption that attacks deviate from normal behavior.

Keywords: Intrusion Detection System (IDS), Anomaly Detection, Principle Component Analysis (PCA), Wavelet Analysis, Self – Organizing map (SOM) and Machine Learning

Introduction

There are various approaches available for intrusion detection system. The most of the commercial intrusion detection systems is largely network-based, and employs signature based intrusion detection methods which are based on human experts' extensive knowledge of known patterns to identify intrusion current tools completely lack the ability to detect attacks that do not fit a pre-defined signature. An important research focus is anomaly detection. Anomaly detection systems try to flag the observed activities that deviate significantly from the established normal usage profiles as anomalies i.e., possible intrusions. This approach can easily detect the novel attacks.

Anomaly Detection

Anomaly Detection or in other words Deviation Detection is the process of localizing the objects that are different from other objects (anomalies). These anomalous objects lie far away from other data points. They have attribute values that deviate significantly from the expected or typical attribute values. Thus, they help in indicating errors. An anomaly is an anomalous object (point) that is sensibly different from other objects (points).In statistic; an outlier is an observation that is numerically distant from the rest of the data. There are various causes of anomalies such as the objects may be of a different type or class, maybe there are errors in the data collection or during the measurement process or may be some natural variation. This anomaly detection method can be applied for intrusion detection to detect the abnormal behavior or deviation from the normal activity. A simplified Anomaly Detection system model is shown in figure 1.

Important Issues in Anomaly Detection

There are various issues that to be handle before using any anomaly detection scheme. Some of them are as follow:

Ruby Dahiya*
IITM, GGSIPU

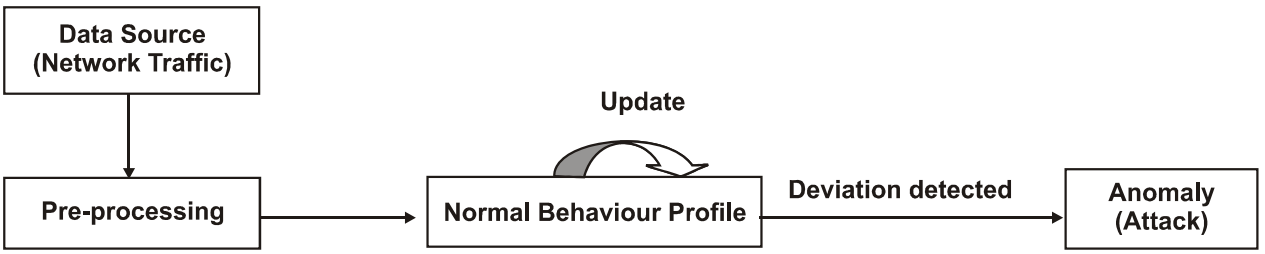

Fig. 1: Simplified Anomaly Detection System

a) *Number of Attributes*: It is very important to work on the no. of attributes and which attributes have to be taken into consideration. Since an object may have many attributes, it may have anomalous values for some attributes or an object may be anomalous even if none of its attribute values are individually anomalous.

b) *Global Vs Local Perspective:* The perspective of the anomalous object may differ from global to local. An object may seem unusual with respect to all objects, but not with respect to its local neighbors.

c) *Degree of Anomaly:* Some objects are more extreme anomalies than others; it's desirable to have some assessment of the degree to which an object is anomalous i.e. it must have some kind of outlier score.

d) *One at Time Vs Many at Once:* It is difficult to decide that which one is better either to remove anomalous objects one at time or to identify a collection of objects together? Two distinct problems: masking, where the presence of an anomaly masks the presence of other; swamping, where normal objects are classified as outliers.

e) *Evaluation:* Another important factor in anomaly detection is to find a good measure of evaluation for the process of anomaly detection when class labels are available and when class labels are not available.

f) *Efficiency:* Calculate the computational cost of the process of anomaly detection scheme priory to make it an efficient system.

Use of Class Labels

a) *Supervised*: In this approach, we need a training set with both anomalous and normal objects.

b) *Unsupervised:* The class labels are not available. We assign a score to each instance (degree of anomaly).

Most of the anomaly – based detection systems are based on supervised approaches [20, 17, 5]. For instance, Audit Data Analysis and Mining (ADAM) [5] employs association rules algorithm in intrusion detection. Actually, one of the most popular ways to undermine anomaly based IDSs is to incorporate some intrusive activities into the training data. The IDSs trained by the training data with intrusive activities will lose the ability to detect this kind of intrusions. Another problem of the supervised anomaly based IDS is high false positive rate when network environment or services are changed. To overcome the limitations of supervised anomaly based systems, a number of IDSs employ unsupervised approaches [6, 18, 9]. Unsupervised anomaly detection does not need attack-free training data. It detects attacks by determining unusual activities from data under two assumptions [9]:

- The majority of activities are normal.
- Attacks statistically deviate from normal activities.

So, this approach has problems with many similar anomalies: can be labeled as normal or have a low outlier score.

There is one more approach known as Semi-supervised. This approach has the training data that contains labeled normal data but has no information

about anomalous objects. We try to find a label or score for anomalous object by using the information from labeled normal objects.

Methods for Anomaly Detection

Statistical Methods

The earliest approach, proposed by Denning, employs statistics to construct a point of reference for system behavior. Statistical methods monitor the user or system behavior by measuring certain variables over time (e.g. login and logout time of each session in intrusion detection domain). The basic models keep averages of these variables and detect whether thresholds are exceeded based on the standard deviation of the variable. More advanced statistical models also compare profiles of long-term and short-term user activities.

For online detection of anomalies processing resource constraints is required. Then some discrete algorithms can be used for the processing of streaming data. In comparison with statistical sampling, streaming peruses every piece of data for the most important information while sampling processes only a small percentage of the data and absorbs all the information therein [25].

The examples of statistical methods are Statistical sequential change-point detection [22,24], Principle Component Analysis [11, 12], Wavelet analysis[2,8] and Covariance Matrix method[26].

Discrete Algorithms

There are two types of discrete algorithms: Heavy Hitters and Heavy-Change. These two are streaming algorithms. In the data stream model, some or all of the input data that are to be operated on are not available for random access from disk or memory, but rather arrive as one or more continuous data streams. For this class of problems, there is a vector $\mathbf{a} = (a_1, \ldots, a_n)$ (initialized to the zero vector 0) that has updates presented to it in a stream. The goal of these algorithms is to compute functions of **a** using considerably less space than it would take to represent **a** precisely. Streams can be denoted as an ordered sequence of points (or "updates") that must be accessed in order and can be read only once or a small number of times.

The Heavy Hitters algorithm is aimed to find those items whose frequencies exceed a threshold during the observation window whereas the goal of heavy-change detection is to efficiently identify the set of flows that have drastic change in traffic volume from one time period to another with small memory requirements and limited state information [26]. Its objective is to find all elements i whose frequency $a_i > T$, say. Some notable algorithms are Count-Min Sketch, Sticky sampling, Sample and Hold, Count-sketch, Sketch-guided sampling etc.

Machine Learning

Machine Learning, a branch of artificial intelligence, was originally employed to develop techniques to enable computers to learn. The core of machine learning deals with representation and generalization. Representation of data instances and functions evaluated on these instances are part of all machine learning systems. Generalization is the property that the system will perform well on unseen data instances. Machine learning approaches attempt to obtain an anomaly detection that adapts to measurements, changing network conditions, and unseen anomalies. Machine learning algorithms can be organized into a taxonomy based on the desired outcome of the algorithm or the type of input available during training the machine.

- The **supervised machine learning** technique finds application in neural networks (NNs) algorithms are trained on labelled examples, i.e., input where the desired output is known. The supervised learning algorithm attempts to generalize a function or mapping from inputs to outputs which can then be used to speculatively generate an output for previously unseen inputs.
- The **unsupervised learning algorithms** operate on unlabelled examples, i.e., input where the desired output is unknown. Here the objective

is to discover structure in the data (e.g. through a cluster analysis), not to generalize a mapping from inputs to outputs. In this mode, networks can learn to pick out structures in their input. One of the most popular models in the unsupervised framework is the self – organizing map (SOM) [19].

- The **transductive inference**, tries to predict new outputs on specific and fixed (test) cases from observed, specific (training) cases.
- The **reinforcement learning** is concerned with how intelligent agents ought to act in an environment to maximize some notion of reward. The agent executes actions which cause the observable state of the environment to change. Through a sequence of actions, the agent attempts to gather knowledge about how the environment responds to its actions, and attempts to synthesize a sequence of actions that maximizes a cumulative reward.

Some Applications Areas of Anomaly Detection

1. *Intrusion detection*: Intrusion detection refers to detection of malicious activity (break-ins, penetrations, and other forms of computer abuse) in a computer related system. An intrusion is different from the normal behavior of the system; and hence anomaly detection techniques are applicable in intrusion detection domain. The key challenge for anomaly detection in this domain is the huge volume of data. Semi-supervised and unsupervised anomaly detection techniques are preferred in this domain.
2. *Fraud detection*: Fraud detection refers to detection of criminal activities occurring in commercial organizations such as banks, credit card companies, insurance agencies, cell phone companies, stock market, etc. The fraud occurs when these users consume the resources provided by the organization in an unauthorized way. The typical approach of anomaly detection techniques is to maintain a usage profile for each customer and monitor the profiles to detect any deviations.
3. *Medical and Public Health Anomaly Detection*: Anomaly detection in the medical and public health domains typically work with patient records. The data can have anomalies due to several reasons such as abnormal patient condition or instrumentation errors or recording errors. Several techniques have also focused on detecting disease outbreaks in a specific area the anomaly detection is a very critical problem in this domain and requires high degree of accuracy.
4. *Industrial Damage Detection*: Industrial units suffer damage due to continuous usage and the normal wear and tear. Such damages need to be detected early to prevent further escalation and losses. The data in this domain is usually referred to as sensor data because it is recorded using different sensors and collected for analysis. Anomaly detection techniques have been extensively applied in this domain to detect such damages.
5. *Image Processing:* Anomaly detection techniques dealing with images are either interested in any changes in an image over time (motion detection) or in regions which appear abnormal on the static image. The anomalies are caused by motion or insertion of foreign object or instrumentation errors. The key challenge in this domain is the large size of the input. When dealing with video data, online anomaly detection techniques are required.

Literature Survey

Based on Statistical Approach

Statistical sequential change-point detection has been used for network anomaly detection.

Thottan et al. [1998] characterized network anomalies with Management Information Base (MIB) variables undergoing abrupt changes in a correlated fashion. Given a set of MIB variables

sampled at a fixed time-interval, the authors compute a network health function by combining the abnormality indicators of each individual MIB variable. This network health function can be used to determine whether there is an anomaly in the network.

Wang et al. [2002] detected SYN flooding attacks based on the dynamics of the differences between the number of SYN and FIN packets, which is modeled as a stationary random process. The better traffic modeling methods that can capture the non-stationary behavior could lead to improved anomaly detection with lower false alarm rates. Also, the accurate characterization of anomalies in terms of abrupt changes in network dynamics is essential for effective anomaly detection.

Principle Component Analysis (PCA, also called Karhunen-Loeve transform) is a coordinate transformation method that maps the measured data onto a new set of axes called Principal Components. Each principal component points in the direction of maximum variation or energy remaining in the data. The principal axes are ordered by the amount of energy in the data they capture.

Lakhina et al.[2004,2005] pioneered the application of Principal Component Analysis (PCA) to network-wide anomaly detection. The basic idea of using PCA for traffic anomaly detection is that: the *k*-subspace obtained through PCA corresponds to the normal behavior of the traffic, whereas the remaining (*n*;*k*) subspace corresponds to either the anomalies or the anomalies and the noise. Each new traffic measurement vector is projected on to the normal subspace and the anomalous subspace. Afterwards, different thresholds can be set to classify the traffic measurement as normal or anomalous.

In [27] **Yin et al. [2004]** proposed a new method takes into account those of frequency property. Since there was no need to consider each system call in each trace or command in each block, the computational cost of the proposed method is low and suitable for real-time intrusion detection. Data found in intrusion detection problem are often high dimensional in nature. By using the proposed method, the high dimensional data can be greatly reduced by projecting them onto a lower dimensional subspace for intrusion detection so that the complexity of the detecting algorithm is significantly reduced.

Experiment results were good in terms of detection accuracy, computational expense and implementation for real-time intrusion detection.

Wavelet analysis is a dimensionality reduction technique. It uses discrete wavelet transform (DWT) which is a linear signal processing similar to discrete Fourier transform (DFT). But it is better lossy compression, localized in space.

Barford et al.[2002] successfully applied wavelet techniques to network traffic anomaly detection. The wavelet analysis in mainly focused on aggregated traffic data in network flows. The authors developed a wavelet system that can effectively isolate both short and long-lived traffic anomalies.

Kim et al.[2008] extended the work in by studying IP packet header data at an egress router through wavelet analysis for traffic anomaly detection. The authors studied the correlation among addresses and port numbers over multiple timescales with discrete wavelet transforms.

Traffic anomalies were detected if historical thresholds were exceeded in the analyzed signal. Wavelet analysis had proved to be an effective anomaly detection method.

In [14] **Wei Lu andAli A. Ghorbani** proposed a completed network anomaly detection approach based on wavelet transformation and the system identification theory. The input signal is a 15-dimensional feature vector, which is defined to characterize the behavior of the network flows. A prediction model for normal daily traffic is established, in which wavelet coefficients play an important role since we use these normal wavelet coefficients as an external input to an ARX model that predicts the approximation coefficient of the signal yet to be seen. The outputs of this traffic

prediction model are called residuals that measure the difference between normal and anomalous activities. The empirical observations show that the peaks of residuals always stand for the location where attacks occur.

Seyed Mahmoud Anisheh and Hamid Hassanpour [2012] worked on an anomaly detection algorithm based on discrete stationary wavelet transform (DSWT) and fractal dimension (FD) [1]. Wavelet technique had been used to exhibit the important underlying unadulterated form of the time series. This pre-processing step is used to increase the accuracy of the proposed method in anomaly detection. The main advantage of DSWT was the preservation of time information of the original signal sequence at each level compared to classical wavelet transform. After applying discrete stationary wavelet transform on the signal representing the network traffic, the fractal dimension of the decomposed signal is calculated in a sliding window. Then, variations of signal fractal dimension are considered for anomaly detection.

A covariance matrix method is a a powerful anomaly detection method to model and detect flooding attacks.

Yeung et al.[2007] developed a covariance matrix method to model and detect flooding attacks. Each element in the covariance matrix corresponds to the correlation between two monitored features at different sample sequences. The norm profile of the normal traffic can then be described by the mathematical expectation of all covariance matrices constructed from samples of the normal class in the training dataset. Anomaly can be detected with threshold-based detection schemes. In [21], **Tavallaee et al. [2008]** the covariance matrix method is extended, where the sign of the covariance matrices is used directly for anomaly detection.

Mandjes et al. [2005] considered anomaly detection in voice over IP network based on the analysis of the variance of byte counts. The authors derived a general formula for the variance of the cumulative traffic over a fixed time interval, which can be used to determine the presence of a load anomaly in the network [15].

Based on Discrete Algorithms

Due to the high link speed and the large size of the Internet, it is usually not scalable to track the per-flow status of traffic. By limiting the number of flows that need to be monitored, sampling can partially solve the scalability problem at the cost of anomaly detection performance. In this area, one important issue is to investigate the tradeoff between the amount of sampled information and the corresponding performance. To address the disadvantages of sampling approaches, there has been extensive research in data streaming algorithms for anomaly detection in high-speed networks. A key difference between streaming and sampling is that streaming peruses every piece of data for the most important information while sampling digests a small percentage of data and absorbs all the information therein.

Estan et al.[2002] initiated a new direction in traffic measurement by recommending concentrating on only large flows, i.e., flows whose volumes are above certain thresholds[7]. The authors also proposed two algorithms for detecting large flows: sample and hold algorithm and multistage filters algorithm.

Cormode et al. [2003, 2004] introduced the Count-Min sketch method to heavy-hitter detection. Sketch is a probabilistic summary data structure based on random projections. The authors noted that it is an open problem to develop extremely simple and practical sketches for data streaming applications [4, 3].

Krishnamurhty et al. [2003] first applied sketch to the heavy-change detection problem. With sketch-based change detection, input data streams are summarized using *k*-ary sketches. After sketches were created, different time series forecast models can be implemented on top of the summaries. Then the forecast errors were used to identify whether there is significant changes in the stream.

The sketch-based techniques uses a small amount of memory and has constant prerecord update and

reconstruction costs, thus it can be used for change detection in high-speed networks with a large number of flows. However, the *k*-ary sketch based change detection has one main drawback: the *k*-ary sketch is irreversible, thus making it impossible to reconstruct the desired set of anomalous keys without querying every IP address or querying every address in the stream if these IP addresses are saved.

Based on Machine Learning

Given a set of training samples, a machine learning view of anomaly detection is to learn a mapping f (:) using the training set, where f (:) is so that a desired performance can be achieved on assigning a new sample x to one of the two categories – normal and anomalous. Specifically, when only normal data is available, learning and thus anomaly detection is unsupervised. When training data and normal data are available, learning/anomaly detection can, be viewed as unsupervised. When anomalous data and normal are both available, learning/anomaly detection becomes supervised, since we have labels or signatures.

Unsupervised Learning examples are clustering, entropy-based, hidden markov model and many more approaches. Learning with additional information becomes supervised learning like probe-measurements.

Kingsly Leung and Christopher Leckie [2005] presented a new density-based and grid-based clustering algorithm that is suitable for unsupervised anomaly detection. They evaluated on the 1999 KDD Cup data set. Their results showed that the accuracy of their approach was close to that of existing techniques and had several advantages in terms of computational complexity.

In [13], **Lima et al. [2010]** proposed a novel model for network anomaly detection combining baseline, K-means clustering and particle swarm optimization (PSO).). The baseline consists of network traffic normal behavior profiles, while K-means is a supervised learning clustering algorithm used to recognize patterns or features in data sets. In order to escape from local optima problem, the K-means is associated to PSO, which is a meta-heuristic whose main characteristics include low computational complexity and small number of input parameters dependence. Anomalous behaviors can be identified by comparing the distance between real traffic and cluster centroids. The obtained detection and false alarm rates are promising.

Wuzuo WANG, Weidong WU [2010] proposed a system model with an explicit algorithm to perform on-line traffic analysis [23]. In this scheme, we first make use of degree distributions to effectively profile traffic features, and then use the entropy to determine and report changes of degree distributions, which changes of entropy values can accurately differentiate a massive network event, normal or anomalous by adaptive threshold. Their scheme found to be accurate and efficient enough to use a little flow header features for capturing fine-grained patterns in traffic distributions. It not only reduces the on-line processing time but increase the detection abilities. The use of entropy can increase the sensitivity of detection to uncover well-known or unknown anomalies and quantify traffic anomalies. An adaptive threshold is also available to lower false alarm rate.

McCallum. A et al.[2005] developed a behavior-based anomaly detection method that detects network anomalies by comparing the current network traffic against a baseline distribution[28]. The packet distribution of the benign traffic was estimated using the Maximum Entropy framework and used as a baseline to detect the anomalies. The method is able to detect anomalies by inspecting only the current traffic instead of a change point detection approach. This is a flexible and fast approach to estimate the baseline distribution, which also gives the network administrator a multi-dimensional view of the network traffic. It also provides information revealing the type of the anomaly detected. It requires a constant memory and a computation time proportional to the traffic rate.

Table 1: The comparative analysis of three approaches for anomaly detection based IDSs

Approach	Data Handling	Data Flow Variation	Nature	Data Knowledge	Online Handling
Statistical Approaches	Yes but Requires large memory space	Difficult to handle	Non-Adaptive	Very Large amount is required	Not Possible
Discrete Algorithms	Yes with no memory space constraint	Easily Handled	Adaptive	Very Small amount is required	Possible
Machine - Learning	Yes with no memory space constraint	Easily Handled	Adaptive	Adequate amount is required	Possible

Result

A comprehensive study of various works done in the field of developing intrusion detection system particularly based on anomaly detection approaches has been made. There are three approaches to detect anomaly in a given network data: statistical, application of discrete algorithms and machine learning. The table-1 depicts the comparison of three approaches on five different parameters.

Conclusion

From the above study, it is very obvious that anomaly detection techniques are very promising in the field of network security. As we discussed that there are some issues at the deployment part which have to be considered. The reasons for these difficulties during implementation may be because of nature of the information that is fed to the anomaly detector could be varied both in format and range and the nature of the anomaly, its frequency of occurrence and resource constraints clearly dictates the detection method of choice. Sampling strategies for multi time scale events with resource constraints is another area where there is a need for improved scientific understanding. Comparing the three approaches, statistical approaches are good when there's sufficient knowledge of the data. The issues related to these approaches: standard or non-standard distribution, single attribute or multivariate data, modeling with a single one or a mixture of distribution. It is tough to capture the statistical dependencies observed in the raw data. To overcome the challenges faced by statistical approaches, streaming algorithms are used. It is also impossible to capture these statistical dependencies unless there are some rule based engines that can correlate or couple queries from multiple streaming algorithms. Machine learning techniques enable the development of anomaly detection algorithms that are non-parametric, adaptive to changes in the characteristics of normal behavior in the relevant network and portable across applications. Machine learning for intrusion detection also faces domain-specific challenges – long training time, poor operational interpretation, the need for outlier detection, very high costs of classification errors. There is a need to investigate the fundamental tradeoffs between the amount/complexity of information available and the detection performance, so that computationally efficient real-time anomaly detection is feasible in practice.

References

1. Anisheh S.M. and Hassanpour H."Designing an Approach for Network Traffic Anomaly Detection" *International Journal of Computer Applications (0975 – 8887)* Volume 37– No.3, January 2012
2. Barford P., Kline J., Plonka D., Ron A., "A Signal Analysis of Network Traffic Anomalies", *Proc. of the 2nd ACM SIGCOMM* , Workshop on Internet Measurements, 71 - 82 (2002)

3. Cormode G., Johnson T., Korn F., Muthukrishnan S., Spatscheck O., Srivastava D. " Holistic UDAFs at Streaming Speeds" *Proc. of ACM SIGMOD*, Paris, France (2004).

4. Cormode G., Korn F., Muthukrishnan S., D. Srivastava D."Finding Hierarchical Heavy Hitters in Data Streams" *Proc. of VLDB*, Berlin, Germany (2003)

5. Daniel Barbarra, Julia Couto, Sushil Jajodia, Leonard Popyack, and Ningning Wu, "ADAM: Detecting Intrusions by Data Mining", *Proceedings of the 2001 IEEE, Workshop on Information Assurance and Security*, T1A3 1100 United States Military Academy, West Point, NY, June 2001.

6. E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, "A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data", *Applications of Data Mining in Computer Security*, Kluwer, 2002.

7. Estan C., Varghese G. "New Directions in Traffic Measurement and Accounting" *Proc. of*

8. Kim S. S., Reddy A., "Statistical Techniques for Detecting Traffic Anomalies Through Packet Header Data", Accepted by IEEE/ACM Tran. Networking , 2008.

9. Kingsly Leung and Christopher Leckie, "Unsupervised Anomaly Detection in Network Intrusion Detection Using Clusters", *Australasian Computer Science Conference*, Newcastle, NSW, Australia, 2005.

10. Krishnamurthy B., Sen S., Zhang Y., Chan Y."Sketch-Based Change Detection: Methods, Evaluation, and Applications", Proc. of ACM SIGCOMM IMC, Florida, USA (2003)

11. Lakhina A., Crovella M, Diot C., " Diagnosing Network-Wide Traffic Anomalies", *Proc. of ACM SIGCOMM* , 2004.

12. Lakhina A., Crovella M., Diot C., "Mining Anomalies Using Traffic Feature Distributions", *Proc. of ACM SIGCOMM*, Philadelphia, PA , 2005.

13. Lima, M.F.; ZarpelaÞo, B.B.; Sampaio, L.D.H.; Rodrigues, J.J.P.C.; AbraÞo, T.; Proença, M.L. "Anomaly detection using baseline and K-means clustering," *Software, Telecommunications and Computer Networks (SoftCOM), 2010 International Conference on* , vol., no., pp.305-309, 23-25 Sept. 2010

14. Lu Wei and Ghorbani Ali A. "Network Anomaly Detection Based on Wavelet Analysis", *EURASIP Journal on Advances in Signal Processing*, Volume 2009, Hindawi Publishing Corporation.

15. Mandjes M., Saniee I., Stolyar A. L., "Load Characterization and Anomaly Detection for Voice over IP traffic*" IEEE Tran. Neural Networks* Vol.16, no.5, 1019-1026 (2005)

16. Nalavade and Meshram, "Intrusion Prevention Systems: Data Mining Approach", *International Conference and Workshop on Emerging Trends in Technology*, ACM 2010.

17. Q.A. Tran, H. Duan, and X. Li, "One-class Support Vector Machine for Anomaly Network Traffic Detection", *The 2nd Network Research Workshop of the 18th APAN*, Cairns, Australia, 2004.

18. Rasheda Smith, Alan Bivens, Mark Embrechts, Chandrika Palagiri, and Boleslaw Szymanski, "Clustering Approaches for Anomaly Based Intrusion Detection", *Walter Lincoln Hawkins Graduate Research Conference 2002 Proceedings*, New York, USA, October 2002.

19. Shah-Hosseini, Hamed; Safabakhsh, Reza (April 2003). "TASOM: A New Time Adaptive Self-Organizing Map", *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics* 33 (2): 271–282.

20. Susan M.Bridges, and Rayford B. Vaughn, "Fuzzy Data Mining and Genetic Algorithms Applied to Intrusion Detection", *Proceedings of the National Information Systems Security Conference (NISSC)*, Baltimore, MD, October, 2000.

21. Tavallaee M., Lu W., Iqbal S. A., Ghorbani A. "A Novel Covariance Matrix Based Approach for Detecting Network Anomalies*" Communication Networks and Services Research Conference* (2008)

22. Thottan M., Ji C., "Proactive Anomaly Detection Using Distributed Intelligent Agents", IEEE *Network.* Vol. 12, no. 5, 21-27, 1998.

23. Wang W., Weidong WU "Online Detection of Network Traffic Anomalies Using Degree Distributions" *Int. J. Communications, Network and System Sciences,* 2010, 3, 177-182

24. Wang, H., Zhang, D., Shin, K. G, "Detecting SYN flooding attacks", *Proc. of IEEE INFOCOM*, 2002.

25. Xu J., "Tutorial on Network Data Streaming", *SIGMETRICS,* 2007.

26. Yeung D. S., Jin S., Wang X., "Covariance-Matrix Modeling and Detecting Various Flooding Attacks", *IEEE Tran. Systems, Man and Cybernetics*, Part A, vol.37, no.2, 157-169 (2007)

27. Yin.F, Wang.J, and Guo C. "A Novel Intrusion Detection Method Based on Principle Component Analysis in Computer Security", *LNCS* 3174, pp. 657–662, 2004 Springer-Verlag Berlin Heidelberg

28. Yu Gu, Andrew McCallum, Don Towsley, "Detecting Anomalies in Network Traffic Using Maximum Entropy Estimation" *Proceedings Of Internet Measurement Conference* 2005, pp. 345–350