Volume 3, No. 1, Jan-Feb 2012

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

CASE STUDY AND REPORT

Available Online at www.ijarcs.info

Survey on Intrusion Detection Approaches

B.Ben Sujitha*
Assistant Professor,
Dept of I T, Noorul Sethu Institute of Technology,
Virudhunagar, Tamilnadu, India
ben_sujitha@rediffmail.com

Dr. V.Kavitha
Director,
University College of Engineering,
Nagercoil, Tamilnadu, India

Abstract: Intrusion detection is a significant focus of research in the security of Computer Systems and Networks. Even though there is an availability of various mechanisms to detect the Intrusion, due to simple lacking the system could not be able to identify the new type of attacks. Probably the Intrusion Detection System may also give false alarm. This paper presents the various approaches in the development of effective Intrusion Detection Systems for computer systems and Distributed computer networks. The importance of this survey is to build the system of a systematic framework which should be a good model to select the best features and the model should be updated dynamically and thus it can be deployed in a robustly. The system should have high accuracy and capable of detecting any type of attacks efficiently.

Keywords: Intrusion Detection System (IDS), Data Mining, Soft Computing, Agent

I. INTRODUCTION

Intrusions are actions that attempt to by pass security mechanisms of computer systems. So they are any set of actions that threatens the integrity, availability, or confidentiality of a network resource. Examples of Intrusion are DOS attack, R2L, U2L, Probe, etc. Intrusion Detection [1] is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of Intrusions, like unauthorized entrance, activity, or file modification [2,3].There are important three steps in the process of Intrusion Detection which are:

a. Monitoring and analyzing traffic.
b. Identifying abnormal activities.
c. Assessing severity and raising alarm.

Intrusion Detection System (IDS) is software that automates the Intrusion Detection process and detects possible Intrusions. Intrusion Detection Systems serve for three essential security functions are monitoring, detecting, and responding to unauthorized activity by company insiders and outsider Intrusion. IDS have been first introduced by James P.Anderson in 1980[4].

Today, Intrusion Detection is one of the high priority and challenging tasks for network administrators and security professionals. Key elements of Intrusion Detection are resources to be protected and Model the behavior of the resources whether they are "normal" or "legitimate" .Efficient methods that compare real-time activities against the models and report probably "intrusive" activities. Denning and Neumann [5] identified four reasons for utilizing Intrusion Detection within a secure computing framework:

a. Many existing systems have security flaws which make them vulnerable, but which are very difficult to identify and eliminate because of technical and economic reasons.
b. Existing system with security flaws cannot be easily replaced by more secure systems because of application and economic considerations.
c. The development of completely secure systems is probably impossible.
d. Even highly secure systems are vulnerable to misuse by legitimate users.

II. CLASSIFICATION

The attacks are classified in the following categories as given in the paper [6]:

a) ***Denial of Service (DOS) Attacks:*** A denial of service attack is a class of attack in which an attacker makes some computing or memory resource too busy or too full to handle valid requests, or denies valid users access to a machine. Examples are Apache2, Back, Land, Mail bomb, SYN Flood, Ping of death, Process table,Smurf, Syslogd, Teardrop, Udpstorm.

b) ***B. Root Attacks:*** User to root exploits are a class of attacks in which an attacker starts out with access to a normal user account on the system and is able to exploit weakness to gain root access to the system.
Examples are Eject, Ffbconfig, Fdformat, Loadmodule, Perl, Ps, Xterm.

c) ***C. Remote to User Attacks****:* A remote to user attack is a class of attacks in which an attacker sends packets to a machine over a network but who does not have an account on that machine; exploits some vulnerability to gain local access as a user of that machine. Examples are Dictionary, Ftp_write, Guest, Imap, Named, Phf,Sendmail, Xlock, Xsnoop.

d) ***D. Probing (Probe):*** Probing is a class of attacks in which an attacker scans a network of computers to gather information or find known vulnerabilities. An attacker with a map of machines and services that are available on

a network can use this information to look for exploits. Examples are Ipsweep, Mscan, Nmap, Saint, Satan.

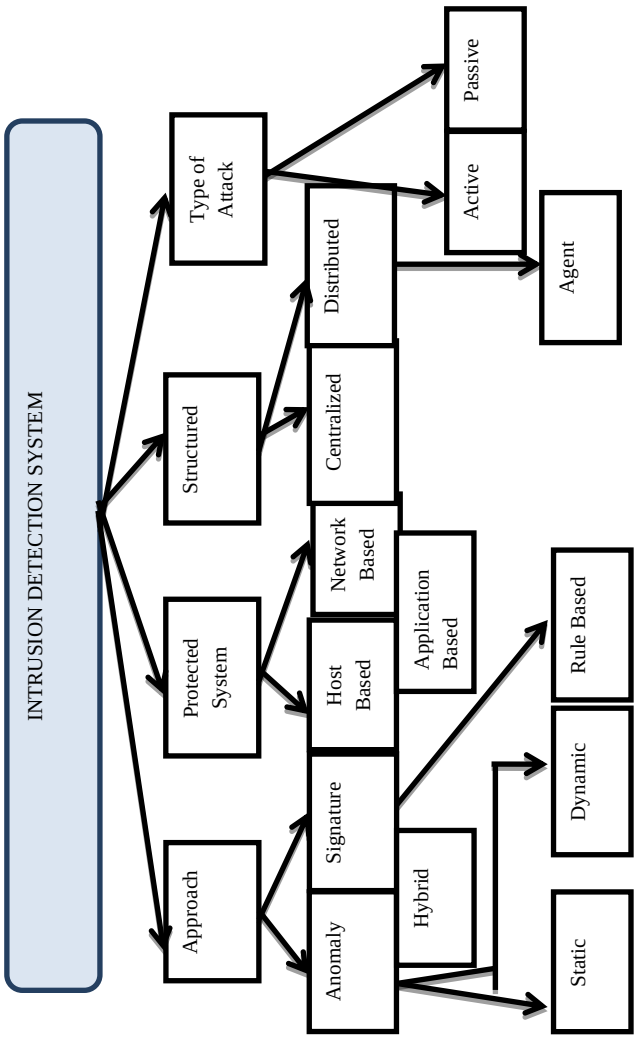

Figure 1: Classification of IDS.

As given in the figure 1 Intrusion Detection systems can be classified into three categories based on types of data[1] they examine. They are Host Based IDS which examine data held on individual computers that serve as hosts. The network architecture of host-based is agent-based, which means that a software agent resides on each of the hosts that will be governed by the system. , Application Based IDS and Network Based IDS which examine data exchanged between computers. More efficient host-based intrusion detection systems are capable of monitoring and collecting system audit trails in real time as well as on a scheduled basis, thus distributing both CPU utilization and network overhead and providing for a flexible means of security administration. Intrusion Detection technique is classified as Misuse Detection, Anomaly Detection and Hybrid Detection.

A. Misuse Detection:

Misuse Detection is the match degree between detection and the unacceptable behavior. The security expert first collects the behavior characteristic of the unusual operation, builds the related characteristic library. This type of detection use well known patterns or signature to identify Intrusion. The system match the record of library, if match exists then it is reported as Intrusion. In this kind of detection, the rate of false alarm is low, but the rate of missing report is high. For the known attack, it may reports the attack type detailed and accurately; but for the unknown attack, it's function is limited. Moreover, the characteristic library must be renewing continually.

Approaches for the misuse detection model are:

- ***a. Expert Systems [7]***, containing a set of rules that describe attacks.
- ***b. Signature Verification [8]***, where attack scenarios are translated into sequences of audit events.
- ***c. Petri Nets [9]***, where known attacks are represented with graphical petri nets.
- ***d. State-Transition Diagrams [10]***, representing attacks with a set of goals and transitions.

The common approach for misuse detection is Signature Verification [8], where a system detects previously seen and known attacks by looking for an invariant signature left by these attacks. This signature is found in audit files, or host-intruded machine, or in sniffers looking for packets inside or outside of the attacked machine.

B. Anomaly Detection:

Anomaly Detection is the deviation between detection and acceptable behavior. It can define each acceptable behavior and unacceptable behavior. The features of Anomaly detection are establishing the normal behavior profiles, Observing and comparing current activities with the (normal) profiles, reporting the significant deviations as Intrusions In this kind of detection, the rate of missing report is low, although it can detect unknown Intrusion effectively, the rate of false alarm is high. In Anomaly detection, the system is Unable to capture sequential interrelation between events.

C. Approaches for Anomaly Detection are:

- ***a. Threshold Detection [12],*** detecting abnormal activity on the server or network, for example abnormal consumption of the CPU for one server, or abnormal saturation of the network.
- ***b. Statistical Measures [13]***, learned from historical values.
- ***c. Rule-Based Measures [14],*** with expert systems.
- ***d. Non-LinearAlgorithms[15],*** such as Neural Networks or Genetic algorithms

The common approach for anomaly detection concerns the statistical analysis, where the user or the system behavior is measured by a number of variables over the time. These variables may be the login and the logout time of each session, the amount of resources consumed during the session, and the resource duration. The major limitation of this approach is to find a correct threshold without frequent false-alarm detection.

III. FEATURE SELECTION

Feature selection and ranking[16]is the important step in classification since the inclusion of irrelevant and redundant features often degrade the performance of classification algorithms both in speed and accuracy. Feature selection is necessary because it is computationally infeasible to use all available features, or because of problems of estimation when limited data samples (but a large number of features) are present. Feature selection from the available data is vital to the effectiveness of the methods employed. Generally the input can be KDD Cup dataset [20], DARPA Data Set or features extracted from the real packet collected. Extracted features can be ranked with respect to their contribution and utilized accordingly.it is very important to consider only the relevant set of features that can best describe the system behavior. This set of features is referred to as axis features.

A minimized set of features is also required to reduce the computational requirements and achieve real time response to Intrusion detection. Hence, selection of appropriate feature set is critical for the performance of the IDS. Lee and Stolfo [17] selected the axis features by empirical knowledge. Mukkamala, Gagnon, and Jajodia [18] discussed three feature reduction methods: significance test, mutual significance test, and rule-based methods. Significance test determines whether or not a feature contributes correctly in classifying an observed user activity as intrusive or non-intrusive. If the significance of a feature is lower than a critical threshold value at a given confidence level, that feature can be eliminated. The mutual significance test determines the inter-feature dependencies. When a feature is present and significant, the other one may be insignificant and thus can be eliminated. The paper [19] reveals the better solution for feature selection with reduced dimensionality.

IV. DATA MINING APPROACHES

Data mining (DM), also called Knowledge-Discovery and Data Mining, is the process of automatically searching large volumes of data for patterns. Data Mining process extracts effective, updated, latent, useful, and the understandable pattern from a lot of incomplete, noise, non-stable, vague and random data. In the Intrusion detection system, the important information comes from the host log, the network data package, the system's log data against applications, and alarm messages. The data mining technology has the huge advantage in the data extracting characteristic and the rule, so it is of great importance to use data mining technology in the Intrusion detection. The data mining technology was first applied in the Intrusion detection research area by Lee and Salvatore J.Stolfo[21]. Data mining methods provide automatic Intrusion detection capabilities. They mine knowledge from audit data to characterize normal and abnormal user behavior. One of the major limitations is that they lack adaptability to changing behavior patterns. As the data mining methods have to deal with large amount of data, a combination of data warehousing and data mining technologies can be utilized for Intrusion detection purposes. Some of the important data mining approaches are discussed.

A. Association Rule and Frequent Episodes:

The goal of mining association rules is to derive multi feature (attribute) correlations from a database table. An association rule is an implication of the form X ->Y [c,s], where X and Y are disjoint itemsets, s is the support of X->Y (indicating the percentage of total records that contain both X and Y), c is the confidence of the rule and is defined as sX->Y/sX [22]. It has been observed that program executions and user activities exhibit frequent correlations among system features. Lee and Stolfo [23] extended the basic association rules algorithms to capture the consistent behaviors in program execution and user activities. The rules mined from audit data are merged and added into an aggregate rule set to form the user's normal profile. Frequent patterns are mined from the sequence of commands during the session and this new pattern set is compared with the normal profile. Similarity functions are used to evaluate deviations involving missing or new rules, violation of the rules (same antecedent but different consequent), and significant changes in support of the rules.

Association rule algorithms find correlations between features or attributes used to describe a data set. The most popular algorithm for mining rules based on two valued attribute is APRIORI algorithm introduced by Agrawal [24]. This algorithm leads to the problem of categorizing numerical attributes. A solution to this problem is given in [25] by transforming quantitative variables into a set of binary variables. This approach has suffered from "sharp boundary ". An alternative solution is given in [26] using fuzzy, offered a smooth transitions from one fuzzy set to another. It is a popular technique, but Association Rule mining usually very slow .It is based on building classifiers by discovering relevant pattern of program and user behavior. It is used to learn the record patterns that describe user behavior. This can deal with symbolic data. Features can be defined in the form of packet and connection details. Mining of features is limited to entry level of the packet. This approach produces large number of records, so it leads to be large number of rules. Thus there is an increase in the complexity of the system.

B. Classification:

Classification Technique is suitable for prediction, since the data have few dimensions equal to the size of the sliding window. The different options for Classification are Decision Trees, SVM, and Naive Bayes. Out of these, Decision Trees provide the best results.

a. Decision Tree:

Decision Trees [11] are structures used to classify data with common attributes. Each decision tree represents a rule which categorizes data according to these attributes. A decision tree consists of *nodes*, *leaves*, and *edges*. A node of a decision tree specifies an attribute by which the data is to be partitioned. Each node has a number of edges which are labeled according to a possible value of the attribute in the parent node. An edge connects either two nodes or a node and a leaf. Leaves are labeled with a decision value for categorization of the data. ID3, C4.5, CART and SPRINT are the best known decision tree algorithms. The paper [27] has suggested ID3 algorithm to construct decision trees from

structured data .The ID3 algorithm uses information theoretic precepts to create efficient decision trees. Given a structured data set, a list of attributes describing each data element and a set of categories to partition the data. The ID3 algorithm determines which attribute most accurately categorizes the data. A node is established and labeled by this attribute.

The edges coming from this node are labeled with the possible values of the partitioning attribute. The C4.5 algorithm [28] which is an upgraded version of ID3 algorithm uses highest Gain ratio for splitting purpose that ensures a larger than average information gain. The C5.0 algorithm [29] improves the performance of building trees using boosting [30] which is an approach to combining different classifiers. But boosting does not always help when the training data contains a lot of noise. Normally C4.5 algorithm is high speed of operation and high attack detection accuracy. When C5.0 performs a classification, each classifier assigns vote and the example is assigned to the class with the most number of votes. CART (Classification and Regression Trees) is a process of generating a binary tree for decision making [31]. CART handles missing data and contains a pruning strategy. The decision tree must be pruned to generalize the information. The SPRINT (Scalable Parallelizable Induction of Decision Trees) Algorithm uses an impurity function called gini index to find the best split[32].

b. Support Vector Machine:

Support vector machines (SVMs)[33] is a supervised learning methods based on statistical learning theory used for classification and regression. They belong to a family of generalized linear classifiers. SVMs attempt to separate data into multiple classes using a hyper-plane. Eskin et al. and Honig et al. [34] used an SVM in addition to their clustering methods for unsupervised learning. The achieved performance is comparable good Mukkamala, Sung, et al. [35] used a more conventional SVM approach. Feature of SVM is good generalization ability of the learning model, which means having small training data set. This approach has good accuracy and has the ability to overcome the curse of dimensionality. SVM constructs the classifier by evaluating an kernel function between two vectors of the training data instead of explicitly mapping the training data into the high dimensional feature space. Therefore SVM is capable of handling large number of features .SVMs are superior to neural nets in both accuracy and speed.

c. Naïve Bayes:

The Naïve Bayes [36] algorithm is a heavily simplified Bayesian probability model. In this model, consider the probability of an end result given several related evidence variables. The probability of end result is encoded in the model along with the probability of the evidence variables .The probability of an evidence variable gives the end result which is assumed to be independent of the probability of other evidence variables. The naïve Bayes classifier operates on a strong independence assumption [37]. This means that the probability of one attribute does not affect the probability of the other. The results of the naïve Bayes classifier are often correct. In the training phase, the Naïve Bayes algorithm calculates the probabilities of a theft given a particular attribute and then stores this probability. This is repeated for each attribute, and the amount of time taken to calculate the relevant probabilities for each attribute. In the testing phase, the amount of time taken to calculate the probability of the given class for each example in the worst case is proportional to n, the number of attributes. Since this approach follows strict independent assumption between the features in an observation, which has the result low at detection accuracy and also the features are correlated.

d. Bayesian Network:

A Bayesian network [38] is a model that encodes probabilistic relationships among variables of interest. This technique is generally used for Intrusion detection in combination with statistical schemes, a procedure that yields several advantages including the capability of encoding interdependencies between variables and of predicting events, as well as the ability to incorporate both prior knowledge and data. This can build a decision network based on special characteristics of individual attacks. a serious disadvantage of using Bayesian networks is that their results are similar to those derived from threshold-based systems, while considerably higher computational effort is required. Although the use of Bayesian networks has proved to be effective in certain situations, the results obtained are highly dependent on the assumptions about the behavior of the target system, and so a deviation in these hypotheses leads to detection errors, attributable to the model considered. The Size of network increases when there is an increase in the number of features.

C. Clustering :

Data clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is useful in Intrusion detection as malicious activity should cluster together, separating itself from non-malicious activity. Clustering provides some significant advantages over the classification techniques already discussed, in that it does not require the use of a labeled data set for training. Frank [39] breaks clustering techniques into five areas: hierarchical, statistical, exemplar, distance, and conceptual clustering, each of which has different ways of determining cluster membership and representation. Clustering was the first choice because the dataset was huge and multidimensional.

a. KMeans:

K Means [40] algorithm uses Normal datasets and clusters the normal behavior points. For the test data set, the probability of its belonging to the most probable cluster is computed. This algorithm can deal with a large number of variables. K-Means may be computationally faster than hierarchical clustering (if K is small). K-Means may produce tighter clusters than hierarchical clustering, especially if the clusters are globular. If it is below a threshold, the instance is flagged as anomalous. This approach did not give very good results. This provide difficulty in comparing quality of the clusters produced (e.g. for different initial partitions or values of K affect outcome). It has Fixed number of clusters can

make it difficult to predict what K should be. It does not work well with non-globular clusters. Different initial partitions can result in different final clusters. It is helpful to rerun the program using the same as well as different K values, to compare the results achieved.

b. KNN Technique for Outlier Detection:

In this method [41] for each test data-point, the distance from its Nearest Neighbor is computed. If this distance is found to be greater than some prespecified threshold, the point is marked as an Outlier. The next step is to establish the threshold. The threshold is obtained by finding out the distance of each training instance from its Nearest Neighbor (or average of distance from k-Nearest Neighbors). The maximum of these distances has been estimated as the value for the threshold.

D. Incremental mining:

The patterns of Intrusions may be dynamic. Intruders may change their strategies over time and the normal system activities may change because of modifications to work practices. Moreover, it is not always possible to predict the level of Intrusions in the future. So it is important that an IDS should have automatic adaptability to new conditions. Otherwise IDS may start to lose its edge. Such adaptability can be achieved by employing incremental mining techniques. Such an adaptive system should use real time data (log of audit records) to constantly update the profile. To make the mining process efficient, the audit data needs to be detailed enough and needs to contain data at the lowest possible level.

Practically, it becomes very difficult to tackle such large data since the whole data set needs to be scanned to compute the support of attributes. Barbara, Jajodia, and Wu [42] described an incremental technique that uses association rules and classification techniques to detect attacks. It does not use the entire data set to mine rules. The rules are categorized according to the time of the day and day of the week. An incremental on-line algorithm is used to detect rules that receive strong support during a sliding window of pre-determined size. It compares these rules with a previously generated profile and selects rules not in that profile. A new rule is reported as suspicious if its support exceeds a threshold. For a set of suspicious rules, a drill-down operation is performed to find the raw data in the audit trail that gave rise to these rules, and these rules are fed to a decision tree that classifies the suspicious activity either as a known or an unknown attack type.

E. Level-wise mining:

Many attacks occur at the application level and can be finished in a single connection. It becomes difficult to detect these by using association rules or frequent episode rules since the support of the generated rule may not be enough to pass the threshold. If the threshold is decreased to that level, many false alarms may be generated. To overcome this, Lee, Stolfo, and Mok [43] proposed a level-wise mining technique that first finds the episodes related to high frequency axis attribute values. Then the support threshold is iteratively reduced to find the episodes related to the low frequency axis values by restricting the participation of the old axis values that already have output episodes.

V. SOFT COMPUTING APROACHES

Soft computing provides tolerance to vagueness. The guiding principle of soft computing is: Exploit the tolerance for imprecision, uncertainty, partial truth, and approximation to achieve tractability, robustness and low solution cost. Soft computing techniques are a natural way of handling the inherent flexibility with which humans communicate, request information, describe events or perform actions. Soft Computing refers to a collection of new computational techniques in computer science, artificial intelligence, machine learning, and many applied and engineering areas where one tries to study, model, and analyze very complex phenomena [44]. Soft computing may be viewed as a foundation component for the emerging field of conceptual intelligence [45]. Soft computing is very useful in Intrusion detection because the techniques used in soft computing give accurate results with high speed which increases the efficiency and performance of the system. The principle constituents of Soft Computing are Fuzzy Logic, Neural Computing, Evolutionary Computation and Machine Learning.

A. Fuzzy Logic:

Fuzzy logic and probabilistic logic are mathematically similar as both have truth values ranging between 0 and 1, but conceptually distinct, due to different interpretation. Fuzzy logic corresponds to “degrees of truth”, while probabilistic logic corresponds to “probability, likelihood”; as these differ, fuzzy logic and probabilistic logic yield different models of the same real-world situations. Fuzzy logic is appropriate for the Intrusion detection problem for two major reasons. First, many quantitative features are involved in Intrusion detection. The second motivation for using fuzzy logic to address the Intrusion detection problem is that security itself includes fuzziness [46]. Given a quantitative measurement, an interval can be used to denote a normal value. Then, any values falling outside the interval will be considered anomalous to the same degree regardless of their distance to the interval. The same applies to values inside the interval, i.e., all will be viewed as normal to the same degree. The use of fuzziness in representing these quantitative features helps to smooth the abrupt separation of normality and abnormality and provides a measure of the degree. Fuzzy logic has proved to be a powerful tool for decision making to handle and manipulate Imprecise and noisy data. The fuzzy logic provides some flexibility to the uncertain problem of Intrusion detection and allows much greater complexity for IDS. Most of the fuzzy IDS require human experts to determine the fuzzy sets and set of fuzzy rules. These tasks are time consuming. However, if the fuzzy rules are automatically generated, less time would be consumed for building a good Intrusion classifier and shortens the development time of building or updating an Intrusion classifier.

B. Neural Computing:

A neural network (NN) [47] in the case of artificial neurons called artificial neural network (ANN) or simulated neural network (SNN) is an interconnected group of natural or artificial neurons that uses a mathematical or computational model for information processing based on a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network. The neural network gains knowledge about the transformation to be performed by iteratively learning from a sufficient training set of samples or input-output training pairs. A well-trained network can perform the transformation correctly and also possess some generalization capability [47]. Multi-layer feed forward ANNs are capable of making multi-class classifications, an ANN is employed to perform the Intrusion detection using the same training and testing sets.

Two types of architecture of Neural Networks can be distinguished:

- ***a. Supervised training algorithms***, where in the learning phase the network learns the desired output for a given input or pattern. The well known architecture of supervised neural network is the Multi-Level Perceptron (MLP) [48]; the MLP is employed for Pattern Recognition problems.
- ***b. Unsupervised training algorithms***, in the learning phase itself the network learns without specifying desired output. Self-Organizing Maps (SOM)[49] are popular unsupervised training algorithms . SOM tries to find a topological mapping from the input space to clusters. SOM are employed for classification problems. A good introduction to Neural Networks is available in [50]. Ensemble SOM are able to identify computer attacks and characterize them appropriately with amazing detection rate and false positive rate under test conditions. This is attractive because of considering the properties of event and its capability of processing large amount of data with low computational overhead. The most important property of a Neural Network is to automatically learn / retrain coefficients in the Neural Network according to data inputs and data outputs. It works well with real network traffic and attacks. Neural Networks have been largely employed with success for complex problems such as Pattern Recognition, hand-written character recognition, Statistical Analysis.

The neural network provides flexibility. A Neural Network would be capable of analyzing the data from the network, even if the data is incomplete or distorted. Similarly, the network would possess the ability to conduct an analysis with data in a non-linear fashion. The inherent speed of Neural Networks is another benefit of this approach. Because the output of a Neural Network is expressed in the form of a probability the Neural Network provides a predictive capability to the detection of instances of misuse.

There are two primary reasons why Neural Networks have not been applied to the problem of misuse detection in the past. The first reason relates to the training requirements of the Neural Network. Because the ability of the Artificial Neural Network to identify indications of an Intrusion is completely dependent on the accurate training of the system, the training data and the training methods that are used are critical. The training routine requires a very large amount of data to ensure that the results are statistically accurate. The training of a Neural Network for misuse detection purposes may require thousands of individual attacks sequences, and this quantity of sensitive information is difficult to obtain.

However, the most significant disadvantage of applying Neural Networks to Intrusion detection is the "black box" nature of the Neural Network. The "Black Box Problem" has overwhelmed Neural Networks in a number of applications [51]. This is an on-going area of Neural Network research. Back propagation neural networks (BPN) with sample-query and attribute-query.[52] provides feasibility and effectiveness. Also the training time is reduced. Additionally, the training results are good. It provides a powerful tool to help supervisors analyze, model and understand the complex attack behavior of electronic crime.

C. Evolutionary Computation:

Evolutionary computation is a subfield of artificial intelligence that involves combinatorial optimization problems .Evolutionary computation uses iterative progress, such as growth or development in a population [53]. This population is then selected in a guided random search using parallel processing to achieve the desired end. Such processes are often inspired by biological mechanisms of evolution. Several variants of evolutionary algorithms have been used for designing Intrusion detection systems. Linear Genetic Programming (LGP) is a variant of the conventional Genetic Programming (GP) technique that acts on linear genomes [54]. MultiExpression Programming (MEP) a chromosome encodes more than one problem solution [55]. The chromosome fitness is usually defined as the fitness of the best expression encoded by that chromosome. Shi [56] used genetic algorithms to automatically select the appropriate set of features and tune it for member hip function for a specific type of Intrusion. His method showed improvements over approaches where the features are selected from empirical evidence. Shi [56] used genetic algorithms to automatically optimize the fuzzy-membership function parameters. Deployment of GAs for Intrusion detection offers number of advantages

a. GAs has multiple offspring, hence they are intrinsically parallel, and they can explore the solution space in multiple directions at once.
b. Due to the parallelism that allows them to implicitly evaluate many schemas at once, GAs are particularly well-suited to solving problems where the space of potential solutions is truly huge - too vast to search exhaustively in any reasonable amount of time, as network data is.
c. System based on GAs can easily be re-trained, which provides the possibility of evolving new rules for Intrusion detection. This property offers the adaptability of a GA-based system.

D. Machine Learning:

Machine learning refers to a system capable of the autonomous acquisition and integration of knowledge [57]. This method learns from experience, analytical observation, and other means, results in a system that can improve its speed and performance. The benefits of applying machine learning techniques to this domain that they eliminate the manual knowledge acquisition phase required by rule-based approaches and provide a generalization of the models for the attack types [58].AI could provide significant benefits to Intrusion detection through data reduction, the ability to analyze a collection of data to identify the most important components, and classification, the process of identifying intruders.

In particular, there are four areas where AI and machine learning could be applied to Intrusion detection systems:

a. By using concept learning, the ability to train a system to classify elements into categories, the Intrusion detection system would have enhanced capabilities to differentiate normal activities from intrusive.
b. Clustering, the partitioning of elements into groups based on a specified criterion could be applied to the effective classification of users, groups, sessions, etc.
c. Predictive learning techniques applied to Intrusion detection would allow the system to develop a temporal model of data and permit the system to learn of intrusive behavior from temporal data and sequences of individual events.
d. The ability to extract relevant features from irrelevant data and the possibility of combining relevant features into functions that identify intrusive events.

E. Expert system:

The use of expert system [59] techniques in Intrusion detection mechanisms is a significant milestone in the development of effective and practical detection-based information security systems. An expert system consists of a set of rules that encode the knowledge of a human "expert". These rules are used by the system to make conclusions about the security-related data from the Intrusion detection system. Expert systems permit the incorporation of an extensive amount of human experience into a computer application that then utilizes that knowledge to identify activities that match the defined characteristics of misuse and attack. Unfortunately, expert systems require frequent updates to remain current. While expert systems offer an enhanced ability to review audit data, the required updates may be ignored or performed infrequently by the administrator. At a minimum, this leads to an expert system with reduced capabilities. At worst, this lack of maintenance will degrade the security of the entire system by causing the system's users to be misled into believing that the system is secure, even as one of the key components becomes increasingly ineffective over time.

VI. AGENTS

An agent is a physical or logical entity characterized by the attributes are Autonomy, Mobility, Rationality, Reactivity, Inferential capability, Pro-activeness, Social ability, Robust behavior and Scalability. The degree of protection against every malicious action is directly related to the time and effort spent in constructing and managing security systems. These actions can be identified at the moment using complex tools to continuously monitor and warn of suspect activities. They provide an attractive and radically different approach to the applications design process. Mobile Agents (MA) [60] is a particular type of software agent, having the capability to move from one host to another. Based on data mining techniques, this agent provides a high accuracy for predicting different behaviors in network computers. Agent provides a high accuracy for predicting different behaviors in network computers. Agent is feasible for detecting attacks within a distributed environment [60]. Consequently, MAD-IDS [61] architecture provides many favorable characteristics, such as high accuracy, good scalability and adaptability. The advantages include: overcoming network latency, reducing network load, executing asynchronously and autonomously, adapting dynamically, operating in heterogeneous environments, and having robust and fault-tolerant behavior. The obvious disadvantage of using MAs is the concern that they will introduce vulnerabilities into the network. However, this is not the only disadvantage to implementing Mobile Agent Intrusion Detection System (MAIDS). The detection agents are lightweight since they can function independently from the heavyweight learning agents, in time and locale, so long as it is already equipped with the rule sets. A detection agent can report new instances of Intrusions by transmitting the audit records to the learning agent, which can in turn compute an updated classifier to detect such Intrusions, and dispatch them to all detection agents.MA solutions may not perform fast enough to meet the IDS's needs. In addition, the MAs may contain large amounts of code thus prohibiting rapid transfers between

VII. CONCLUSION

Since IDS tools are becoming the need for the day and for security not only in the corporate world but also for network users. Security incidents are becoming more and more common and measures have to be taken to curb such incidents. .In this paper the overview of Intrusion detection techniques and methods are given. It's very important to identify the false alarm rates as low as possible; the false negative alarms should be the minimum to ensure the security of the system. To overcome this limitation, IDS must be capable of adapting to the changing conditions typical of an Intrusion detection environment. From the survey done, the IDS can be developed as a Agent which can be deployed robustly using Genetic Approach in Combining with Neuro Fuzzy Data mining Approach. The new System can be a Fast and flexible detection System to identify the vast variety of clever and unusual attacks.

VIII. REFERENCES

[1]. V.Mehta,C.Bartzis,H.Zhu,E.M.Clarke and J.Wing ,"Ranking attack graphs" presented at the International Symposium on Recent Advances in Intrusion Detection" , Hambay, Germany ,sep 20-22 , 2006.

[2]. NorthCutt, S. "Network Intrusion Detection: An Analyst's Handbook",New Riders, Indianapolis, 1999.

[3]. http://csrc.nist.gov/publications/nistpubs/800- 31/sp800- 31.pdf

[4]. James P.Anderson ," Computer Security Threat monitoring and Surveillance " Technical report 98, Washington , Pennysylvannia , USA, April 1980.

[5]. Neumann, P.G. "Audit Trail Analysis and Usage Collection and Processing", Technical Report Project 5910, SRI International.

[6]. Lunt, T.,"Detecting Intruders in Computer Systems", In Proceedings of IEEE conference on Auditing and Computer Technology, February 1999.

[7]. Teresa F. Lunt, Ann Tamaru, Fred Gilham,, R.Jagannathan, Caveh Jalali, Peter G. Neumann," A Real Time Intrusion Detection Expert System , Final Technical Report February 28, 1992.

[8]. M. El-Ghali,W. Masri , "Intrusion Detection Using Signatures Extracted from Execution Profiles", Proceedings of the ICSE Workshop on Software Engineering for Secure Systems, IEEE Computer Society Washington, DC, USA 2009.

[9]. Chien-Chuan Lin,Ming-ShiWang," Depth Evaluation Intrusion Detection using Coloured Stochastic Petri Nets",Intelligence and Security Informatics,IEEE International Conference on 17-20, pp 248 – 250, June 2008.

[10]. Porras, P.A. Kemmerer,"CA Penetration state transition analysis: A rule-based Intrusion Detection Approach", Computer Security Applications Conference Proceedings, pp 220-229, Dec 1992.

[11]. Dewan Md. Farid, Nouria Harbi, Emna Bahri, Mohammad Zahidur Rahman, ,"Classification in Adaptive Intrusion Detection using Decision Tree",World Academy of Science, Engineering and Technology, 2010.

[12]. Sathish Alampalayam P. Kumar, Anup Kumar, S.Srinivasan,"Statistical Based Intrusion Detection Framework using Six Sigma Technique",IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.10, October 2007

[13]. Denning, D. E., "An Intrusion-detection model", IEEE Transactions on Software Engineering 13 (2), 222-232, February, 1987.

[14]. P.A. Bonatti, J.L. De Coi, D. Olmedilla, and L. Sauro,"A Rule-Based Trust Negotiation System", IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 11, November 2010.

[15]. Muna Elsadig Mohamed Brahim Belhaouari SamirAzween Abdullah," Immune Multiagent System for Network Intrusion Detection using Non-linear Classification Algorithm", International Journal of Computer Applications ,Vol 12, No.7, Dec 2010.

[16]. Sung A. H. "Ranking Importance of Input Parameters of Neural Networks," Expert Systems with Applications, Vol. 15, pp.405-41,1998.

[17]. W. Lee, S. J. Stolfo, and K. W. Mok, "Adaptive Intrusion detection: A Data Mining Approach," Artificial Intelligence Review, vol. 14, no. 6, pp.533–567, 2000.

[18]. Mukkamala, S. and A. H. Sung, "Identifying Significant Features for Network Forensic Analysis using Artificial Intelligent Techniques", International Journal of Digital Evidence 1 (4), 1-17, 2003.

[19]. Ahmad, Azween B Abdulah, "Feature Subset Selection for Network Intrusion Detection Mechanism Using Genetic Eigen Vectors".International Conference on Telecommunication Technology and Applications,Proc .of CSIT vol.5, 2011.

[20]. h//www.kdd.ics.uci.edu/databases/kddcup99/task.html , Knowledge discovery in databases DARPA archive.

[21]. W. Lee and S. Stolfo, "Data Mining Approaches for Intrusion Detection", Proc. 7th USENIX Security Symposium , San Antonio, TX, 1998, 79-94,1998.

[22]. R. Agrawal, T. Imielinski, and A. Swami, "Mining Association Rules Between Sets of Items in Large Databases", Proc. ACM SIGMOD International Conf. on Management of Data, Washingtion, DC,pp. 207-216,1993.

[23]. Wenke Lee and Salvatore J. Stolfo, "Data Mining Approaches for Intrusion Detection//', In Proceedings of the 1996 IEEE Symposium on Security and Privacy, pages 120-128, Los Alamitos, CA, 1996.

[24]. R.Agrawal , R.Srikent , " Fast Algorithm for mining Association rules " , In Proc. of the 20th International Conference on very large database held in sanho, China, Sep 12-14, 1994.

[25]. S. Bridges and R. Vaughn, "Fuzzy Data Mining and Genetic Algorithms Applied to Intrusion Detection", Proc. 23rd National Information Systems Security Conf., Baltimore, MA, 2000.

[26]. Shingo Mabu, Ci Chen, Nannan Lu, Shimada, K,Hirasawa, K,"An Intrusion-Detection Model Based on Fuzzy Class-Association-Rule Mining Using Genetic Network Programming", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol 41, issue 1pp. 130-139, Jan.2011.

[27]. Shekhar R. Gaddam, Vir V. Phoha ,"Clustering and ID3 Decision Tree Learning Methods" IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 3, Mar 2007.

[28]. J.R.Quilon,"C4.5 Programs for Machine Learning ",Margon Kaufmann publishers , San Makeo , CA 1993.

[29]. Rasha G Mohammed and Awad M Awadelkarim, "Design and Implementation of a Data Mining Based Network Intrusion Detection Scheme",Asian Journal of Infromation Technology, pp 136-141,2011.

[30]. Y.Freund, R Schapre," A Decision Theoretic Generation of On line Learning and an Application to Boosting", Journal of Computer and System Science , vol 55 ,pp 119-189, 1999.

[31]. L.Breiman ,J H Friedman, RA Olshan and CJ Sfone ," Classification and Regression Trees ", Statistics Publishng series , Wardsworth , Belmont, 1984.

[32]. John Shafer, Rakesh Agrawal and Manish Mehta ,"SPRINT: A Scalable Parallel Classifier for Data Mining" ,Proc. of the Very Large Data Base Conference , Bombay , India , Sep 1996.

[33]. Snehal A. MulayP.R. DevaleG.V. Garje," Intrusion Detection System using Support Vector Machine and Decision Tree", International Journal of Computer Applications Vol 3, No.3, June 2010.

[34]. Eskin, E., A. Arnold, M. Preraua, L. Portnoy, and S. J. Stolfo, "A Geometric Framework for UnSupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data", In D.

Barbar and S. Jajodia (Eds.), Data Mining for Security Applications. Boston: Kluwer Academic Publishers, May 2002.

[35]. Snehal A. MulayP.R. DevaleG.V. Garje," Intrusion Detection System using Support Vector Machine and Decision Tree", International Journal of Computer Applications Vol 3, No.3, June 2010.

[36]. M. Panda and M.R.Patra,"Network Intrusion Detection using Naïve Bayes" International Journal of Computer Science and Network Security, Vol. 7, No. 12, pp. 258-263, 2007.

[37]. N.B. Amor, S. Benferhat, and Z. Elouedi, "Naive Bayes vs.Decision Trees in Intrusion Detection Systems," Proc. ACM Symp.Applied Computing (SAC '04), pp. 420-424, 2004.

[38]. C. Kruegel, D. Mutz, W. Robertson, and F. Valeur, "Bayesian Event Classification for Intrusion Detection," Proc. 19th Ann.Computer Security Applications Conf. (ACSAC '03), pp. 14-23, 2003.

[39]. Frank, J., "Artificial Intelligence and Intrusion Detection: Current and Future Directions", In Proc. of the 17th National Computer Security Conference, Baltimore, MD. National Institute of Standards and Technology (NIST), 1994.

[40]. Y.Guan,A,A Ghorbani and N.Balacel," Kmeans : A Clustering Method for Intrusion Detection", Proc. In IEEE Can.Electr.Computing Engineering ,vol 2 ,pp 1083-1086,2003.

[41]. S.Bay and M.Schwabacher ," Mining Distance Based Outliers in Near Linear Time with Randomization and a Simple Pruning Rule",Proc. inACM SIGKDD Int.Conf.Knowl.Discov,Data Mining , pp 29-38,2003.

[42]. D.Barbara, S.Jajodia and N.Wu, " Mining Unexpected Rules in Network Audit Trails ", Personal communication 2000.

[43]. W.Lee, S.Stolfo and K.Mok," Mining Audit Data to Build Intrusion Detection Models" , Proc. 4th Int. Conf. on Knowledge Discovery and Data Mining , New York city , Ny , pp 66-72,1998.

[44]. Adel Nadjaran Toosi a,Mohsen Kahani," A New Approach to Intrusion Detection Based on an Evolutionary Soft Computing Model using Neuro-Fuzzy Classifiers" , Elsevier Journal ,Computer Communications vol .30,pp. 2201–2212,2007.

[45]. Surat Srinoy, Werasak Kurutach, Witcha Chimphlee, and Siriporn Chimphlee," Network Anomaly Detection using Soft Computing", World Academy of Science, Engineering and Technology, 2005.

[46]. S.T Sarasamma , Q A Zhu and J Huff ," Hierrarchiacl Kohonenen Net for Anomaly Detection in Network Security",IEEE Transaction on Systems, Man and Cybernatics part B , Cybernatics , 35(2), pp. 302-312,2005.

[47]. L.Delooze ,"Attack Characterization and Intrusion Detection using an Ensamble of Self Organizing Maps", Proceedings of the 2006 IEEE workshop on Infromation Assurance ", United states Military Academy , West Point , NY ,2006.

[48]. Horeis T., "Intrusion Detection with Neural Networks - Combination of Self-organizing Maps and Radial Basis Function Networks for Human Expert Integration", Tech. Rep., University of Passau, Department of Mathematics and Computer Science, Germany, 2003.

[49]. Ray-I Chang, Liang-Bin Lai , Wen-De Su, Jen-Chieh Wang, Jen-Shiang Kouh, "Intrusion Detection by Backpropagation Neural Networks with Sample-Query and Attribute-Query" ,International Journal of Computational Intelligence Research,ISSN 0973-1873 Vol.3, No. 1 , pp. 6-10,2007.

[50]. Brameier. M. and Banzhaf. W., "A Comparison of Linear Genetic Programming and Neural Networks in Medical Data Mining", IEEE Transactions on Evolutionary Computation, Vol: 5(1), pp. 17-26, 2001.

[51]. G.V.S.N.R.V.Prasad Y.Dhanalakshmi Dr.V.Vijaya Kumar Dr I.Ramesh Babu," Modeling An Intrusion Detection System Using Data Mining And Genetic Algorithms Based On Fuzzy Logic", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.7, July 2008.

[52]. YU Yan, and Huang Hao, "An Ensemble Approach to Intrusion Detection Based on Improved Multi-Objective Genetic Algorithm" ,Journal of Software, vol. 18, no. 6, June pp. 1369-1378,2007.

[53]. F. Shi, Genetic Algorithms for Feature Selection in an Intrusion Detection Application" , Masters Thesis, Mississippi State University, Mississippi State, MS, 2000.

[54]. Horeis T., "Intrusion Detection with Neural Networks - Combination of Self-organizing Maps and Radial Basis Function Networks for Human Expert Integration", Tech. Rep., University of Passau, Department of Mathematics and Computer Science, Germany, 2003.

[55]. Cohen W., "Learning Trees and Rules with Set-Valued Features", American Association for Artificial Intelligence (AAAI), 1996.

[56]. Debar, H., Becker, M., and Siboni, D., "A Neural Network Component for an Intrusion Detection System", IEEE Computer Society Symposium on Research in Security and Privacy, Los Alamitos, CA, pp. 240–250, Oakland, CA, May 1992.

[57]. http://www.cerias.purdue.edu/research, "Autonomous Agents for Intrusion Detection".

[58]. F. A. Barika1 & N. El Kadhi,"MA IDS : Mobile Agents for Intrusion Detection System" ,IEEE International Advance Computing Conference (IACC 2009)

