ISSN: 2456-2955
Volume 2, Number 2
UNIVERSAL JOURNAL OF COMPUTERS & TECHNOLOGY

Intrusion Detection Techniques

A.Vikramajit, N. Adithya khanna
Department of Computers and Technology
Sri Manakula Vinayagar Engineering Colleg, India

ABSTRACT

Network security has been one of the most important problems in Computer Network Management and Intrusion is the most publicized threats to security. In recent years, intrusion detection has emerged as an important field for network security. IDSs obtain better results when each class of attacks is treated as a separate problem and handled by specialized algorithms. Now in days various model and method are available for intrusion detection. In this paper, we present a study of intrusion detection. Detection method to improve the detection rate & helping the users to develop secure information systems.

Keywords

Clustering, data mining, intrusion detection system, network security.

I. INTRODUCTION

Intrusion detection systems (IDSs) are monitoring devices that have been added to the wall of security in order to prevent malicious activity on a system. This work focuses on network intrusion detection systems (NIDSs) mainly because they can detect the widest range of attacks compared to other types of IDSs. Network IDSs analyze traffic to detect on-going and incoming attacks on a network. Nowadays, commercial IDSs mainly use a database of rules, called signatures, to try to detect attacks on a network or on a host computer. Intrusion detection systems are monitoring devices which are used to detect intrusions on a computer or a network. Intrusions are unauthorized and anomalous activities which were defined by Christopher Kruegel et al. as \ a sequence of related actions performed by a malicious adversary that results in the compromise of a target system" [1]. An intrusion detection system is an indispensable tool for network administrators because without such a device, it would be impossible to analyze the huge amount of packets traversing current networks every second. After more than thirty years of intensive research on intrusion detection systems, the field is still open to further investigations especially regarding the accuracy of the detection. Moreover, variants of known attacks as well as new attacks can often go through the system without being detected.

The goals of the IDS provide the requirements for the IDS Policy. Potential goals include:

- Detection of attacks
- Prevention of attacks
- Detection of policy violations
- Enforcement of use policies
- Enforcement of connection policies
- Collection of evidence

Following are the factors for the measurement of IDS.

• False positive (FP): Or false alarm, Corresponds to the number of detected attacks but it is in fact normal.

• False negative (FN): Corresponds to the number of detected normal instances but it is actually attack, in other words these attacks are the target of intrusion detection systems.

• True positive (TP): Corresponds to the number of detected attacks and it is in fact attack.

• True negative (TN): Corresponds to the number of detected normal instances and it is actually normal.

II. TYPES OF INTRUSION DETECTION SYSTEMS

Accordance with analytical methods, the Intrusion Detection System can be divided into two categories, one is Abnormal Detection, and the other is Misuse Detection or Signature Detection.

A. Anomaly Detection

Anomaly systems adopt the opposite approach, which is, to know what is normal, and then find the deviations from the normal behavior. These deviations are considered as anomalies or possible intrusions. Anomaly detection systems rely on knowledge of normal behavior to detect any attacks. Thus attacks, including new ones are detected as long as the attack behavior deviates sufficiently from the normal behavior. However, if the attack is similar to the normal behavior, it may not be detected. Moreover, it is difficult to associate deviations with specific attacks. As the users change their behavior, normal behavior should be redefined..

B. Misuse detection

Misuse detection systems use a priori knowledge on attacks to look for attack traces. In other words, they detect intrusions by knowing what the misuse is [2]. Signature (rule) based systems are the most common examples of the misuse detection systems. In signature based detection, attack signatures are sought in the monitored resource. Signature based systems, by definition, are very accurate on known attacks which are included in their signature database. Moreover, since signatures are associated with specific misuse behavior, it is easy to determine the attack type. However, their detection capabilities are limited to those within signature database. As the new attacks are discovered, a signature database requires continuous updating to include the new attack signatures.

The goal of intrusion detection is to build a system which would automatically scan network activity and detect such attacks. Once an attack is detected, the system administrator could be informed and thus take corrective action.

Generally, there are four categories of attacks . They are:

1. DoS (Denial of Service) – trying to prevent a legitimate user from accessing the service in the target machine.

2. Probe – scanning a target machine for information about potential vulnerabilities.

3. R2L (Remote to Local) – when attacker attempts to obtain non-authorized access into a machine or network.

4. U2R (User to Root) – when target machine is already invaded, but the attacker attempts to gain access with super-user privileges.

III. SURVEY OF INTRUSION DETECTION SYSTEM USING DIFFERENT APPROACHES

A. SURVEY ON INTRUSION DETECTION METHODS [14]

Intrusion Detection methods like pattern matching, state full pattern matching, protocol decode-based analysis etc and how fuzzy clustering can apply in IDS.

There are different methods for Intrusion Detection. Some of the methods adopted are the following:

1) Pattern Matching:

Pattern matching is based on looking for a fixed sequence of bytes in a single packet. As its name suggests, it is an approach that is fairly rigid but simple to employ. In most cases the pattern is matched against only if the suspect packet is associated with a particular service or, more precisely, destined to/from a particular port [3, 4]. The structure of a signature based on the simple pattern-matching approach might be as follows if the packet is IPv4 and TCP and the destination port is 2222 and the payload contains the string "foo," fire an alarm. This example of a pattern match, of course, is a very simple one, but the variations from this point are also Simplistic.

2) State Full Pattern Matching:

A more sophisticated method is state full pattern matching-based analysis. This method of signature development adds to the pattern match the concept that because a network stream comprises more than single atomic packets, matches should be made in context within the state of the stream. This means that systems that perform this type of signature analysis must consider arrival order of packets in a TCP stream and should handle matching patterns across packet boundaries [5].

3) Protocol Decode-Based Analysis:

To Protocol decode-based signatures are in many ways intelligent extensions to state full pattern matches. This class of signature is implemented by decoding the various elements in the same manner as the client or server in the conversation would. When the elements of the protocol are identified, the IDS apply rules defined by the RFCs to look for violations. In some instances, these violations are found with pattern matches within a specific protocol field, and some require more advanced techniques that account for such variables as the length of a field or the number of arguments.

4) FUZZY CLUSTERING FOR IDS:

To The underline premise of our intrusion detection model is to describe attacks as instances of an ontology using a semantically rich language like DAML. This ontology capture information attacks such as the system component it affects, the consequences the attacks the mean of attack the location of attacker. Such target - centric ontology has been developed by under conferral, hence our intrusion detection model consist of two phases. The initial phase's data mining techniques to analyze data stream that capture process, system and network states and detect anomalous behavior and the second or high level phase reasons over data that is representative of the anomaly defined as instance of ontology. One way to build the models from these data streams is to use fuzzy clustering in which dissimilar matrix of object to be cluster as input. The objective function are based on selecting, representative objects from the features set in such a way that total fuzzy dissimilarity within each cluster is minimized [6, 7].

B. Intrusion Survey on Data Mining Techniques for Intrusion detection[15]

Data mining is used to clean, classify and examine large amount of network data. Since a large volume of network traffic that requires processing, we use data mining techniques. Different Data Mining techniques such as clustering, classification and association rules are proving to be useful for analyzing network traffic.

Authors presented a four step approach for the Generalized working of IDS

- Data collection: - It involves collecting network traffic using particular software and thus helps to get the information about the traffic like types of packets, hosts and protocol details.
- Feature Selection: - The collected data is substantially large because of the huge network traffic; we generate feature vectors that contain only necessary information. In network-based intrusion detection, it can be IP header information, which consists of source and destination IP address, packet type, layer 4 protocol type and other flags.
- Analysis: - The collected data is analyzed in this step to determine whether data is anomalous or not. Here we use various methods for detecting intrusions.
- Action :- IDS alarm the system administrator that an attack has happened and it tells about the nature of the attack.IDS also participate in controlling the attacks by closing the network port or killing the processes

Data Mining is used in variety of applications that requires data analysis. Now a day's data mining techniques plays an important role in intrusion detection systems. Different data mining techniques like Classification, Clustering and Association rules are frequently used to acquire information about intrusions by observing network data.

1) Classification

Classification method can be useful for both misuse detection and anomaly detection, but it is more commonly used for misuse detection. Authors presented a data classification for intrusion detection that can be achieved by the following steps:-

1. In order to study about the classification models of the normal and abnormal sequences of system calls, we want to supply it with a training data set, containing pre-labeled normal or abnormal sequences. Different techniques like linear discrimination, decision tree or rule based methods is used to scan the network traces. Then generate a collection of unique sequence of system calls and named it as normal list.

2. Next scan each of the intrusion traces. Find each sequence of system calls in the normal list. If an exact match can be found then labeled it as normal. Otherwise it is labeled as abnormal.

3. Next ensure that the normal traces consist of all possible normal short sequence of system calls. An intrusion trace contains combination of normal and abnormal sequences of system calls since abnormal sequence only appear in some places.

2) Clustering

Clustering can be applied on both Anomaly detection and Misuse detection. Authors presents basic steps involved in identifying intrusion are follows:-

1. Find the largest cluster, which consists of maximum number of instances, and label it as normal.

2. Sort the remaining clusters in an ascending order of their distances to the largest cluster.

3. Select the first K1 clusters so that the number of data instances in these clusters sum up to ¼`N, and label them as normal, where ` is the percentage of normal instances.

4. Label all other clusters as malicious.

5. After clustering, heuristics are used to automatically label each cluster as either normal or malicious. The self labeled clusters are then used to detect attacks in a separate test dataset.

3) Association Rule

Association rule mining in intrusion detection is very useful in many ways. An author presents basic steps for incorporating association rule for intrusion detection as follows:-

1. First network data need to be arranged into a database table where each row is an audit record and each column is a field of the audit records.

2. It is always shows that the intrusions and user activities shows frequent correlations among network data. Consistent behaviors' in the network data can be captured in association rules.

3. Also rules based on network data can continuously merge the rules from a new run to the aggregate rule set of all previous runs.

4. Thus with the association rule, we get the capability to capture behavior in association rule for correctly detecting intrusions and hence lowering the false alarm rate.

C. Design of Intrusion Detection System Based on Data Mining Algorithm [16]

Internet technology has developed rapidly and both software system and hardware equipment have improved greatly in recent years. However, internet brings people not only convenience but also great potential threats. Facts show that potential safety hazards exist from the emergence of internet. As a kind of effective information security safeguard measure, intrusion detection makes up for the defects of traditional security protection techniques. As a kind of effective data analysis method, data mining is introduced into intrusion detection systems. This puts forward the idea of applying data mining technology to intrusion detection systems and then designs data preprocessing module, association analysis module and cluster module respectively.

Data mining can find potential and useful knowledge from a mass of data. The advantages of applying data mining to an intrusion detection system lie in that -

1) The system can produce an accurate detection model from a mass of audit data automatically to reduce artificial intervention and it can be used to construct an intrusion detection system in various computing environments because of mechanical ness and universality of mining process itself.

2) In recent years, the rapid development of data mining technology has got a large quantity of algorithms from the fields, such as statistics, pattern recognition, machine learning and database, etc, and some algorithms are particularly useful for intrusion detection, such as classification analysis, cluster analysis, association rule analysis and sequential pattern analysis, etc, the previous studies show that applying these technologies to intrusion detection is feasible and effective.

D. Intrusion Detection System Based on Improved K-means Clustering Algorithm [17]

K-means algorithm's shortcomings about dependence and complexity, puts forward an improved clustering algorithm through studying on traditional means clustering algorithm. The new algorithm learns the strong points from the k-medoids and improved relations trilateral triangle theorem. The experiments proved that the new algorithm could improve accuracy of data classification and detection efficiency significantly. The results show that this algorithm achieves the desired objectives with a high detection rate and high efficiency.

1) The concept of intrusion detection is proposed by James P. Anderson[8] more early in 1980, he defines the intrusion is a kind of visit information or operation information which is latent, deliberate and unauthorized that makes the system untrustworthy or can't be used. As an important link of information security safeguard, the intrusion detection can perform well on atonement the problem that can't be solved by the protection mechanism such as the firewall, access control, security route and so on.

2) In order to reduce the dependence of k-means algorithm on the starting value selection, Kaufman and Rouseeuw [9] have used k-medoids algorithm, which replace the mean value of the data object with the nearest point to the cluster. In order to solve that the difference between the selection of central point of the k-medoids algorithm and the fact is too large, a k-medoids cyclic method is also proposed. It uses one kind of dynamic central point renewal strategy making the criterion function drop quickly. But this method is taking sacrificing the certain time complexity as the price, which is excessively high when the data quantity is large.

E. Intrusion Detection Method Based on Improved DBSCAN [18]

DBSCAN is an efficient method of clustering analysis algorithm in data mining. In this method, a model of intrusion detection system is built, and an improved DBSCAN algorithm IIDBG is applied to detection engine, according to the property of DBSCAN. In IIDBG, distance calculation formula and clusters merger process are

improved based on DBSCAN and IDBC. Experiment results prove this method reduces the false negative rate, and promotes the performance of intrusion detection system, comparing DBSCAN, IDBC.

1) DBSCAN (Density-Based Spatial Clustering of Application with Noise) has been concerned by some researchers, main two advantages: first, it is not sensitive to the order of data input; second, it can find clusters of arbitrary shape in spatial database with noise.

2) In reference [10], the author use traditional DBSCAN algorithm for intrusion detection, the density high enough area generates a cluster, and merge some small clusters. This paper proposes a more reasonable density-based clustering algorithm for intrusion detection (IIDBG), using rational method for calculating the distance and the merger process is optimized reasonably based on [8], and design the method of select parameter selection..

3) When DBSCAN algorithm is used to process intrusion data, found that many small clusters generated after cluster through the experiment results, and part of small clusters contains number of normal records more than 70%, resulting in a high false alarm rate. Moreover, these small clusters are closer to normal clusters. In order to solve this problem, reference [10] proposed IDBC (Improved Density Based Clustering Algorithm) to merge some small clusters. In merging, through calculation the distance of boundary points, to represents the distance between two clusters.

4) When density-based clustering algorithm is applied to intrusion detection, there are some drawbacks. After detailed analyze DBSCAN and IDBC in [10], this paper proposed improved density-based clustering algorithm IIDBG to improve these drawbacks, with using more rational method for calculating the distance and cluster merging process.

F. Intrusion Detection based on K-Means Clustering and OneR Classification [19]

The approach, KM+1R, combines the k-means clustering with the OneR classification technique. The KDD Cup ’99 set is used as a simulation dataset. The result shows that our proposed approach achieve a better accuracy and detection rate, particularly in reducing the false alarm.

Related work and research publications based on hybrid approaches have been widely explored such as in [11]. The detection rate (DR), false positive (FP), false negative (FN), true positive (TP), false alarm (FA), and accuracy for each approach are also investigated. Each approach has distinctive strengths and weakness. Some approaches possess strength in detection but high in false alarm and vice versa. For instance, in [13] the author proposed a new three-level decision tree classification, which focuses on the detection rate. Authors [12] model the IDS using a hierarchical hybrid intelligent system with the combination of decision tree and support vector machine (DT-SVM). While DT-SVM produces high detection rate, it lacks in the ability to differentiate attacks from normal behavior. More recently, approach as suggested by author [11] offers a high detection rate but comes with high false alarm rate as compared to others. In short, a number of hybrid techniques have been proposed in intrusion detection fields and related work; but there are still room to improve the accuracy and detection rate as well as the false alarm rate

The main goal to utilize K-Means clustering approach is to split and to group data into normal and attack instances. K-Means clustering methods partition the input dataset into k- clusters according to an initial value known as the seed points into each cluster’s centroids or cluster centers. The mean value of numerical data contained within each cluster is called centroids.

The K-Means algorithm works as follows:

1. Select initial centers of the K clusters. Repeat step 2 through 3 until the cluster membership stabilizes.

2. Generate a new partition by assigning each data to its closest cluster centers.

3. Compute new clusters as the centroids of the clusters.

One-R algorithm choose attribute with lowest error rate as its “one rule”. A proportion of instances that do not belong to the majority class of the corresponding attribute value will contribute to the error rate. The OneR algorithm works as follows:

1. From clustered set, create a rule set for each value of each attribute predictor as in step i, ii, iii and iv. i. Count how often each value of target class appears, ii. Find the most frequent class, iii. Make a rule set assign that class to this value of attribute predictor, iv. Calculate the total error occurs in the rules set for each attribute predictor.

2. Pick the best attribute predictors which have a smallest total error and this as a classification rules.

G. A Graph –based clustering algorithm for anomaly intrusion detection[20]

A graph based approach is proposed by Zhou Mingqiang et al [20] they proposed graph-based intrusion detection algorithm by using outlier detection method that based on local deviation coefficient (LDCGB). Compared to other intrusion detection algorithm of clustering, this algorithm is unnecessary to initial cluster number. Meanwhile, it is robust in the outlier’s affection and able to detect any shape of cluster rather that the circle one only. Moreover, it still has stable rate of detection on unknown or muted attacks. Table 1 show comparision between different techniques .

Table 1. show comparision

s. no.	Author	Detection Rate	Accuracy	False positive rate	Falsenegative rate
1	Sanoop Mallissery et. al.	Yes	Yes	No	No
2	Deepthy K Denatious & 2Anita John	Yes	Yes	No	No
3	Changxin Song, Ke Ma	Yes	Yes	No	No
4	Li Tian1, Wang Jianwen1	Yes	Yes	Yes	No
5	Li Xue-yong, Gao Guo	Yes	Yes	Yes	No
6	Z. Muda, W. Yassin	Yes	Yes	Yes	No
7	Zhou Mingqiang et. al.	Yes	Yes	No	No

IV. CONCLUSION

Network security situation awareness system and intrusion detection is a new research domain, and it has great importance in improving abilities of responding to emergences, reducing losses of network attacks, revealing abnormally intrusions, enhancing system abilities of fighting back. The paper gives the study of network security situation awareness model and intrusion detection model.

V. FUTURE WORK

In future we can design hybrid model combine with the concept of fuzzy c-means & probabilistic neural network. Shows that the fuzzy C-means & PNN is an effective method for achieve a better accuracy, detection rate, failure analysis rate & false alarm.

REFERENCES

[1] Christopher Kruegel, Fredrik Valeur, and Giovanni Vigna. “Intrusion Detection and Correlation: Challenges and Solutions”, volume 14 of Advances in Information Security. Springer- Verlag, 2005

[2] R. A. Kemmerer and G. Vigna, “ Intrusion detection: A brief history and overview”,IEEE Security and Privacy, April 2002

[3] Sandeep Kumar, Eugene H.Spat'tord, "An Application of Pattern Matching in Intrusion Detection", Technical Report 94-0 13, Department of © 20 1 1 lET 228 Computer Sciences, Purdue University, West Lafayette, March 2004.

[4] Sandeep Kumar, Eugene H.Spat'tord, "A Pattern Matching Model for Misuse Intrusion Detection", The COAST Project, Department of Computer Sciences, Purdue University, West Lafayette, IN 47907- 1 398.

[5] Gonzalo Navarro, Mathieu Raffinot, "Flexible Pattern Matching in Strings", Cambridge University Press 2002, ISBN 0-52 1-8 13077.

[6] John E. Dickerson, Jukka Juslin, Ourania Koukousoula, Julie A. Dickerson, "Fuzzy Intrusion Detection", Electrical and Computer Engineering Department, Iowa State University, Ames, IA, USA.

[7] K.e.C.Chan and W.H.Au, "Mining Fuzzy Association Rules", Proc.Of ACM CIKM, 1997, pp.209- 2 15.

[8] Spector, James P Anderson, “Computer Security Threat Monitoring and Surveillance”, Fort Washington, PA, 1980.

[9] JIAO Licheng Jiao, Fang Liu, and Jing Liu, “Intelligent Data Mining and Knowledge Discovery”, Xi’an: Xidian University Press,2006,pp.325-326.

[10] Yang Jian, “An Improved Intrusion Detection Algorithm Based on DBSCAN”, Micro Computer Information, 25,1008-0570(2009)01- 3-0058-03, 58-60,2009.

[11] C.F. Tsai, C.Y. Lin, “A triangle area based nearest neighbors approach to intrusion detection”, Pattern Recognition 43(1) 222-229, 2010.

[12] S.Peddabachigari,A.Abraham,c.Grosan,J.Thomas, “modeling intrusion detection system using hybrid intelligent system”,Journal of Network and Computer Applications 30 114-132, 2007.

[13] C. Xiang, P.C. Yong, L.S. Meng, “Design of multiple level hybrid classifierfor intrusion detection system using bayesian clustering and decision tree”, Pattern Recognition Letters 29 918-924,2008.

[14] Sanoop Mallissery , Jeevan Prabhu,and Raghavendra Ganiga, “Survey on intrusion detection method”,2011.

[15] Deepthy K Denatious & 2Anita John, “Survey on Data Mining Techniques to Enhance Intrusion Detection”,2012.

[16] Changxin Song, Ke Ma Institute of Computer Information & Technology of Qinghai Normal University Network Center of Qinghai Normal University Qinghai, China., “Design of Intrusion Detection System Based on Data Mining Algorithm”,2009.

[17] Li Tian1, Wang Jianwen1 Department of Computer Science, North China Electric Power University (NCEPU), Baoding 071003, China, “Research on Network Intrusion Detection System Based on Improved K-means Clustering Algorithm”,2009

[18] Li Xue-yong, Gao Guo- “A New Intrusion Detection Method Based on Improved DBSCAN”,2010.

[19] Z. Muda, W. Yassin, M.N. Sulaiman and N.I.Udzir, “Intrusion Detection based on K-Means Clustering andOneR Classification”,2011. Zhou Mingqiang,HuangHui,WangQian, “A Graph-based Clustering Algorithm for Anomaly Intrusion Detection”, 2012