Singaporean Journal Scientific Research (SJSR)
ISSN: 2231 - 0061Vol.4, No.1pp. 221 - 224

available at: : http://www.iaaet.org/sjsr

JPCAP, WINPCAP Used For Network Intrusion Detection System

[1]Kapil Kumar Nagwanshi ,[2]Susanta Kumar Satpathy and [3]Ruchi Jain
[1,2,3]RungtaCollege of Engg& Technology, Bhilai. Chhattisgarh, India
[1]kapilkn@gmail.com, [2]sks_sarita@yahoo.com, [3]cool_ruchijain@rediffmail.com

ABSTRACTIntrusion detection systems serve three essential security functions: they monitor, detect, and respond to unauthorized activity by company insiders and outsider intrusion. Intrusion detection systems use policies to define certain events that, if detected will issue an alert. In other words, if a particular event is considered to constitute a security incident, an alert will be issued if that event is detected. Certain intrusion detection systems have the capability of sending out alerts, so that the administrator of the IDS will receive a notification of a possible security incident in the form of a page, email, or SNMP trap. Many intrusion detection systems not only recognize a particular incident and issue an appropriate alert, they also respond automatically to the event. Such a response might include logging off a user, disabling a user account, and launching of scripts.Of the security incidents that occur on a network, the vast majority (up to 85 percent by many estimates) come from inside the network. These attacks may consist of otherwise authorized users who are disgruntled employees. The remainder comes from the outside, in the form of denial of service attacks or attempts to penetrate a network infrastructure. Intrusion detection systems remain the only proactive means of detecting and responding to threats that stem from both inside and outside a corporate network. As stated previously, intrusion detection is the process of monitoring computers or networks for unauthorized entrance, activity, or file modification. IDS can also be used to monitor network traffic, thereby detecting if a system is being targeted by a network attack such as a denial of service attack. There are two basic types of intrusion detection: host-based and network-based. Each has a distinct approach to monitoring and securing data, and each has distinct advantages and disadvantages. In short, host-based IDSs examine data held on individual computers that serve as hosts, while network-based IDSs examine data exchanged between computers

***Keywords*—** Data mining, intrusion, intrusion detection, network security, clustering.

1. INTRODUCTION

Number of hacking and intrusion incidents is increasing alarmingly each year as new Technology rolls out. Unfortunately in today's digitally connected world, there is no place to hide. DNS, Snooker, Newsgroups, web site trawling, e-mail properties etc. are just some of the many ways in which you can be found. In this research work, it has designed and build an Intrusion Detection System (IDS) that implement pre-defined algorithms for identifying the attacks over a network. The Java programming language is used to develop the system, JPCap must be used to provide access to the win cap. The packets in the network are captured online i.e., as they come on the interface of the network. The IDS is designed to provide the basic detection techniques so as to secure the systems present in the networks that are directly or indirectly connected to the Internet.

Deliberate attempt to A enter a network and break the security of the network and thus breaking the confidentiality of the information present in the systems of the network. The person who tries to attempt such an action is called as an Intruder and the action can be termed as Network Intrusion. The network administrator is supposed to protect his network from such persons and this software can help his in his efforts.

An Intrusion Detection System (IDS) is a system that is responsible for detecting anomalous, inappropriate, or other data that may be considered unauthorized occurring on a network. An IDS captures and inspects all traffic, regardless of whether it's permitted or not. Based on the contents, at either the IP or application level, an alert is generated.

The security of a computer system is compromised when an intrusion takes place. An intrusion can be defined as ``any set of actions that attempt to compromise the integrity, confidentiality or availability of a resource". Intrusion prevention techniques, such as user authentication (e.g. using passwords or biometrics), avoiding programming errors, and information protection (e.g., encryption) have been used to protect computer systems as a first line of defense. Intrusion prevention alone is not sufficient because as systems become ever more complex, there are always exploitable weakness in the systems due to design and programming errors, or various ``socially engineered" penetration techniques. For example, after it was first reported many years ago, exploitable ``buffer overflow" still exists in some recent system software due to programming errors. The policies that balance convenience versus strict control of a system and information access also make it impossible for an operational system to be completely secure.

Use definitions from the pioneering work in intrusion detection.

- **Risk**: Accidental or unpredictable exposure of information, or violation of operations integrity due

to the malfunction of hardware or incomplete or incorrect software design.

- **Vulnerability**: A known or suspected flaw in the hardware or software or operation of a system that exposes the system to penetration or its information to accidental disclosure.
- **Attack**: A specific formulation or execution of a plan to carry out a threat.
- **Penetration**: A successful attack -- the ability to obtain unauthorized (undetected) access to files and programs or the control state of a computer system.

2. CLASSIFICATION

Intrusion detection techniques are traditionally categorized into two methodologies:

A. Anomaly detection

Anomaly detection is based on the normal behaviour of a subject (e.g., a user or a system); any action that significantly deviates from the normal behaviour is considered intrusive. Anomaly detection techniques assume that all intrusive activities are necessarily anomalous. This means that if it could establish a "normal activity profile" for a system, this could, in theory, flag all system states varying from the established profile by statistically significant amounts as intrusion attempts. However, if it is considered that the set of intrusive activities only intersects the set of anomalous activities instead of being exactly the same, it will find a couple of interesting possibilities: (1). Anomalous activities that are not intrusive are flagged as intrusive. (2). Intrusive activities that are not anomalous result in false negatives (events are not flagged intrusive, though they actually are). This is a dangerous problem, and is far more serious than the problem of false positives.

B. Misuse Detection

Misuse detection catches intrusions in terms of the characteristics of known attacks or system vulnerabilities, any action that conforms to the pattern of a known attack or vulnerability is considered intrusive. The concept behind misuse detection schemes is that there are ways to represent attacks in the form of a pattern or a signature so that even variations of the same attack can be detected. This means that these systems are not unlike virus detection systems, they can detect many or all known attack patterns.

Alternatively, IDSs may be classified into host-based IDSs, distributed IDSs, and network-based IDSs according to the sources of the audit information used by each IDS. Host-based IDSs get audit data from host audit trails and usually aim at detecting attacks against a single host; distributed IDSs gather audit data from multiple hosts and possibly the network that connects the hosts, aiming at detecting attacks involving multiple hosts. Network-based IDSs use network traffic as the audit data source, relieving the burden on the hosts that usually provide normal computing services.

3. DATA MINING BASICS

The term data mining is frequently used to designate the process of extracting useful information from large databases. Similarly the term knowledge discovery in databases (KDD) is used to denote the process of extracting useful knowledge from large data sets. Data mining, by contrast it refers to one particular step in this process. Specifically, the data mining step applies so-called data mining techniques to extract patterns from the data. Additionally, it is preceded and followed by other KDD steps, which ensure that the extracted patterns actually correspond to useful knowledge. Indeed, without these additional KDD steps, there is a high risk of finding meaningless or uninteresting patterns (Fayyad, 1998; Klemettinen et al., 1997; Stedman, 1997). In other words, the KDD process uses data mining techniques along with any required pre- and post-processing to extract high-level knowledge from low-level data. In practice, the KDD process is interactive and iterative, involving numerous steps with many decisions being made by the user.

4. PROBLEM DESCRIPTION

A major shortcoming of current IDSs that employ data mining methods is that they can give a series of false alarms in cases of a noticeable systems environment modification. There can be two types of false alarms in classifying system activities in case of any deviation from normal patterns: false positives and false negatives. False positive alarms are issued when normal behaviours are incorrectly identified as abnormal and false negative alarms are issued when abnormal behaviours are incorrectly identified as normal.

Though it's important to keep both types of 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, an IDS must be capable of adapting to the changing conditions typical of an intrusion detection environment. For example, in an academic environment, the behaviour patterns at the beginning of a semester may be different than the behaviour patterns at the middle/end of the semester.

If the system builds its profile based on the audit data gathered during the early days of the semester, then the system may give a series of false alarms at the later stages of the semester. System security administrators can tune the IDS by adjusting the profile, but it may require frequent human intervention. Since normal system activities may change because of modifications to work practices, it is important that an IDS should have automatic adaptability to new conditions. Otherwise, an 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.

One straightforward approach can be to regenerate the user profile with the new audit data. But this would not be a computationally feasible approach. Each of these deviations can represent an intrusion or a change in behaviour. In case of a change in system behaviours, the base profile must be updated with the corresponding change so that it does not give any false positives alarms in future. This means that the

system needs a mechanism fordeciding whether to make a change or reject it. If the system tries to make a change to the base profile every time it sees a deviation, there is a potential danger of incorporating intrusive activities into the profile.

5. SYSTEM DESCRIPTION

The central theme of present approach is to apply data mining techniques K-Means clustering for intrusion detection in network. Data mining generally refers to the process of (automatically) extracting models from large stores of data. The recent rapid development in data mining has made available a wide variety of algorithms, drawn from the fields of statistics, pattern recognition, machine learning, and database. Introducing the concept of intrusion detection in 1980, defined an intrusion attempt or a threat to be the potential possibility of a deliberate unauthorized attempt to access information,

- manipulate information, or
- Render a system unreliable or unusable.

Since then, several techniques for detecting intrusions have been studied. This paper discusses why intrusion detection systems are needed, the main techniques, present research in the field, and possible future directions of research. In the following sections, it use definitions from the pioneering work in intrusion detection.

A. k-Means Clustering

K- Means is an iterative clustering algorithm in which items are moved among set of clusters until the desired set is reached. A high degree of similarity among elements in the clusters is obtained, while a high degree of dissimilarity among elements in different clusters is achieved simultaneously.

This algorithm assumes that the desire number of clusters, K, is an input parameter. The initial values for the means are arbitrarily assigned. These could be assigned randomly or perhaps could use the values from the first K input items themselves.

The cluster mean of k= $\{t_{i1},t_{i2},\ldots\ldots t_{im}\}$ is defined as

$\mathbf{m_i = 1/m \sum_{ij}}$

B. k-Means Clustering Algorithm

1. Input:
2. $D=\{t_1,t_{2,\ldots\ldots\ldots\ldots\ldots},t_n\}$ // Set of elements
3. K // Number of desire clusters
4. Output:
5. S // Set of clusters
6. K-means algorithm:
7. Assign initial values for means m_1,m_2 ,……..,$m_{K;}$
8. Repeat
9. Assign each item t_i to the cluster which has the closest mean;
10. calculate new mean for each cluster;
11. Until convergence criteria is met;
12.

c. Winpcap:

WinPcap is an open source library for packet capture and network analysis for the Win32 platforms. Most networking applications access the network through widely used operating system primitives such as sockets. It is easy to access data on the network with this approach since the operating system copes with the low level details (protocol handling, packet reassembly, etc.) and provides a familiar interface that is similar to the one used to read and write files.Sometimes, however, the 'easy way' is not up to the task, since some applications require direct access to packets on the network. That is, they need access to the "raw" data on the network without the interposition of protocol processing by the operating system.

d.Jpcap:

Jpcap is an open source library for capturing and sending network packets from Java applications. It provides facilities to:

- Capture raw packets live from the wire.
- Save captured packets to an offline file, and read captured packets from an offline file
- Filter the packets according to user-specified rules before dispatching them to the application.
- send raw packets to the network Jpcap is based on libpcap/winpcap, and is implemented in C and Java.

6. CONCLUSION

This paper has proposed a systemic framework that employs data mining techniques for intrusion detection. This framework consists of classification, clustering and frequency episodes programs, which can be used to (automatically) construct detection models. The experiments on log data and network data demonstrated the effectiveness of clustering models in detecting anomalies. The accuracy of the detection models depends on sufficient training data and the right feature set. The k-means clustering algorithm can be used to compute the consistent patterns from audit data. These frequent patterns form an abstract summary of an audit trail, and therefore can be used to: guide the audit data gathering process; provide help for feature selection; and discover patterns of intrusions. This is the initial stages of present research; much remains to be done including the following tasks:

- Implement a support environment for system builders to iteratively drive the integrated process of pattern discovering, system feature selection, and construction and evaluation of detection models.
- Investigate the methods and benefits of combining multiple simple detection models. It is needed to use multiple audit data streams for experiments.
- Implement a prototype agent-based intrusion detection system.
- Evaluate present approach using extensive audit data sets.

REFERENCES

[1] Fayyad, U., “Mining Databases: Towards Algorithms for Knowledge Discovery” Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 1999.

[2] S.J. Stolfo, S. Hershkop, K. Wang, O. Nimeskern, and C.W. Hu, Behavior Profiling of Email, First NSF/NIJ, ISI, 2003;

[3] W. Lee, S.J. Stolfo, K.W. Mok, Algorithms for Mining System Audit Data, in Proc. KDD, 1999;

[4] W.W. Cohen, Fast Effective Rule Induction, in 12th Conference on Machine Learning, CA, 1995;

[5] W. Lee, S. Stolfo, Data Mining Approaches for Intrusion Detection, in 7th Usenix Security, 1998.