Optimizing Network Security with a Hybrid Intrusion Detection System

Gabriel Santos
Department of Computer Engineering, University of São Paulo, Brazil

ABSTRACT

Intrusions detection systems (IDSs) are systems that try to detect attacks as they occur or after the attacks took place. IDSs collect network traffic information from some point on the network or computer system and then use this information to secure the network. Intrusion detection systems can be misuse-detection or anomaly detection based. Misuse-detection based IDSs can only detect known attacks whereas anomaly detection based IDSs can also detect new attacks by using heuristic methods. In this paper we propose a hybrid IDS by combining the two approaches in one system. The hybrid IDS is obtained by combining packet header anomaly detection (PHAD) and network traffic anomaly detection (NETAD) which are anomaly-based IDSs with the misuse-based IDS Snort which is an open-source project. The hybrid IDS obtained is evaluated using the MIT Lincoln Laboratories network traffic data (IDEVAL) as a testbed. Evaluation compares the number of attacks detected by misusebased IDS on its own, with the hybrid IDS obtained combining anomaly-based and misusebased IDSs and shows that the hybrid IDS is a more powerful system..

I. INTRODUCTION

Nowadays with the spreading of the Internet and online procedures requesting a secure channel, it has become an inevitable requirement to provide the network security. There are various threat sources including software bugs mostly as the operating systems and software used becomes more functional and larger in size. Intruders who do not have rights to access these data can steal valuable and private information belonging to network users. Firewalls are hardware or software systems placed in between two or more computer networks to stop the committed attacks, by isolating these networks using the rules and policies determined for them. It is very clear that firewalls are not enough to secure a network completely because the attacks committed from outside of the network are stopped whereas inside attacks are not. This is the situation where intrusions detection systems (IDSs) are in charge. IDSs are used in order to stop attacks, recover from them with the minimum loss or analyze the security problems so that they are not repeated [1]. IDSs collect information from a computer or a computer network in order to detect attacks and misuses of the system. Many IDSs only analyze the attacks and some of them try stopping the attack at the time of the intrusion. Three types of data are used by IDSs. These are network traffic data, system level test data and system status files [2,3]. In ‘‘2003CSI/FBI Computer Crime and Security Survey” it has been stated that the IDS usage in 1999 had been 42% and this ratio has become 73% in year 2003. This great improvement shows that IDSs are very important as security technologies. This paper is organized as follows: intrusion detection systems are described in Section 2, IDS types are explained in Section 3: Snort is the chosen system as misuse-based IDS; PHAD and NETAD are chosen as anomaly-based IDSs. Section 4 gives a brief description of the hybrid IDS we propose in this paper. The newly obtained hybrid IDS is evaluated in Section 5 and finally Section 6 includes conclusion.

II. HISTORY

The goal of intrusion detection is to monitor the network assets to detect anomalous behavior and misuse in network [16]. Intrusion detection concept was introduced in early 1980’s after the evolution of internet with surveillance end monitoring the threat [17]. There was a sudden rise in reputation and incorporation in security infrastructure. Since then, several events in IDS technology have advanced intrusion detection to its current state [16]. James Anderson's wrote a paper for a government organization and imported an approach that audit trails contained important information that could be valuable in tracking misuse and understanding of user behavior [16]. Then the detection appeared and audit data and its importance led to terrific improvements in the subsystems of every operating system [16]. IDS and Host Based Intrusion Detection System (HIDS) were first defined. In 1983, SRI International and Dorothy Denning began working on a government project that launched a new effort into intrusion detection system development [17]. Around 1990s the revenues are generated and intrusion detection market has been raised. Real secure is an intrusion detection network developed by ISS. After a year, Cisco recognized the priority for network intrusion detection and purchased the Wheel Group for attaining the security solutions [17]. The government actions like Federal Intrusion Detection Networks (FID Net) were designed under Presidential Decision Directive 63 is also adding impulse to the IDS [17]

III. IDS TYPES

There are two approaches to analyzing of events using IDSs. These are misuse-based and anomaly-based approaches. Misuse-based IDSs aim to distinguish events that violate system policy. Anomaly-based IDSs try analyzing abnormal activities and flag these activities as attacks. Both approaches have advantages and disadvantages when compared to each other [1,2,5]. Snort is the most commonly used signature-based intrusion detection system. Snort is a network intrusion detection system that runs over IP networks analyzing real-time traffic for detection of misuses [6]. Snort depends on a template-matching scheme and makes content analysis. It has the ability to flag alerts depending on pre-defined misuse rules and saves packets in tcpdump files or in plain text files. Snort is preferred to be used in academic research projects as it is an opensource tool and for this reason we have also chosen Snort as the signature-based intrusion detection system in our work. Anomaly detection based intrusion detection systems are separated into many sub-categories in the literature including statistical methodologies [7–10], data mining [11,12], artificial neural networks [13], genetic algorithms [14] and immune systems [15,16]. Among these sub-categories, statistical methods are the most commonly used ones in order to detect intrusions by analyzing abnormal activities occurring in the network. PHAD [17] and NETAD [18] statistical methods are chosen as the anomaly-based intrusion detection systems in this paper. We have implemented a hybrid IDS by mounting anomalybased IDSs PHAD and NETAD to Snort as a preprocessor. PHAD is different than the other conventional network-based anomaly detection systems for two reasons. First, it models protocols rather than user behaviors. Second, it uses a time-based model depending on the rapid change of network statistics in short term. PHAD flags only the first anomaly it detected as an alert even if there is a series of the same anomaly recurring. This feature of PHAD helps reducing the number of false alerts. NETAD, models single packets like PHAD, uses dynamic-conditioned rules like ALAD [19], and rule verification like LERAD [20]. Its greatest contribution is modeling values that are not new. 3.1. Misuse-based IDSs Misuse detectors analyze system activities and try to find a match between these activities and known attacks having definitions or signatures introduced to the system beforehand [1,2,21]. Advantages: Misuse detectors are very efficient in detecting attacks without signaling false alarms (FA). Misuse detectors can quickly detect specially designed intrusion tools and techniques. Misuse detectors provide systems administrators an easy to use tool to monitor their systems even if they are not security experts. Disadvantages: Misuse detectors can only detect attacks known beforehand. For this reason the systems must be updated with newly discovered attack signatures. Misuse detectors are designed to detect attacks that have signatures introduced to the system only. When a well-known attack is changed slightly and a variant of that attack is obtained, the detector is unable to detect this variant of the same attack.

Principles and assumptions

in Intrusion Detection Denning defines the principle for characterizing a system under attack. The principle states that for a system which is not under attack, the following three conditions hold true: 1. Actions of users conform to statistically predictable patterns. 2. Actions of users do not include sequences which violate the security policy. 3. Actions of every process correspond to a set of specifications which describe what the process is allowed to do. Systems under attack do not meet at least one of the three conditions. Further, intrusion detection is based upon some assumptions which are true regardless of the approach adopted by the intrusion detection system. These assumptions are: 1. There exists a security policy which defines the normal and (or) the abnormal usage of every resource. 2. The patterns generated during the abnormal system usage are different from the patterns generated during the normal usage of the system; i.e., the abnormal and normal usage of a system results in different system behavior. This difference in behavior can be used to detect intrusions. As we shall discuss later, different methods can be used to detect intrusions which make a number of assumptions that are specific only to the particular method. Hence, in addition to the definition of the security policy and the access patterns which are used in the learning phase of the detector, the attack detection capability of an intrusion detection system also depends upon the assumptions made by individual methods for intrusion detection.

Components of intrusion detection systems

An intrusion detection system typically consists of three sub systems or components: 1. Data Preprocessor – Data preprocessor is responsible for collecting and providing the audit data (in a specified form) that will be used by the next component (analyzer) to make a decision. Data preprocessor is, thus, concerned with collecting the data from the desired source and converting it into a format that is comprehensible by the analyzer.Data used for detecting intrusions range from user access patterns (for example, the sequence of commands issued at the terminal and the resources requested) to network packet level features (such as the source and destination IP addresses, type of packets and rate of occurrence of packets) to application and system level behavior (such as the sequence of system calls generated by a process.) We refer to this data as the audit patterns. 2. Analyzer (Intrusion Detector) – The analyzer or the intrusion detector is the core component which analyzes the audit patterns to detect attacks. This is a critical component and one of the most researched. Various pattern matching,

machine learning, data mining and statistical techniques can be used as intrusion detectors. The capability of the analyzer to detect an attack often determines the strength of the overall system. 3. Response Engine – The response engine controls the reaction mechanism and determines how to respond when the analyzer detects an attack. The system may decide either to raise an alert without taking any action against the source or may decide to block the source for a predefined period of time. Such an action depends upon the predefined security policy of the network The authors define the Common Intrusion Detection Framework (CIDF) which recognizes a common architecture for intrusion detection systems. The CIDF defines four components that are common to any intrusion detection system. The four components are; Event generators (E-boxes), event Analyzers (A-boxes), event Databases (D-boxes) and the Response units (R-boxes). The additional component, called the D-boxes, is optional and can be used for later analysis.

IV. PROPOSED WORK

We use two classification techniques for our proposed architecture, in a combined manner. Consequently, an increasing number of approaches have been developed for accomplishing such purpose, including k-nearest-neighbor (KNN) classification, Naïve Bayes classification, support vector machines (SVM), decision tree (DT), neural network (NN), and maximum entropy. Our choice among all available classification techniques is depends upon our studies about all classifier. We put our motivations for these classifiers in below topic at a glance.

V. FUNCTIONS OF IDS

The IDS consist of four key functions namely, data collection, feature selection, analysis and action, which is given in Figure 3. International Journal of Security, Privacy and Trust Management (IJSPTM) Vol 4. Functionality of IDS

Data collection This module passes the data as input to the IDS. The data is recorded into a file and then it is analyzed. Network based IDS collects and alters the data packets and in host based IDS collects details like usage of the disk and processes of the system.
Feature Selection To select the particular feature large data is available in the network and they are usually evaluated for intrusion. For example, the Internet Protocol (IP) address of the source and target system, protocol type, header length and size could be taken as a key for intrusion [15].
Analysis The data is analyzed to find the correctness. Rule based IDS analyze the data where the incoming traffic is checked against predefined signature or pattern [15]. Another method is anomaly based IDS where the system behavior is studied and mathematical models are employed to it [15].
Action It defines about the attack and reaction of the system. It can either inform the system administrator with all the required data through email/alarm icons or it can play an active part in the system by dropping packets so that it does not enter the system or close the ports [15].

VI. TOOLS IN INTRUSION DETECTION

An intrusion detection product available today addresses a range of organizational security goals .This section discusses about the security tools.

SNORT Snort is lightweight and open source software. Snort uses a flexible rule-based language to describe the traffic [6].From an IP address; it records the packet in human readable form. Through protocol analysis, content searching, and various pre-processors Snort detects thousands of worms, vulnerability exploit attempts, port scans, and other suspicious behavior.
OSSEC-HIDS OSSEC (open source security) is free open source software. It will run on major operating system and uses a Client/Server based architecture. OSSEC has the ability to send OS logs to the server for analysis and storage. It is used in powerful log analysis engine, ISPs, universities and data centres. Authentication logs, firewalls are monitored and analysed by HIDS.
FRAGROUTE It is termed as fragmenting router. Here, from the attacker to the fragrouter the IP packet is sent and they are then fragmented and transformed to the party.
HONEYD Honeyd is a tool that creates virtual hosts on the network [6]. The services are used by the host Honeyd allows a single host to request multiple addresses on a LAN for networks simulation. It is possible to knock the virtual machines or to trace route them [6]. Any type of service on the virtual machine can be simulated according to a simple configuration file [6].
KISMET It is a guideline for WIDS (Wireless intrusion detection system).WIDS compromises with packet payload and happenings of WIDS. It will find the burglar access point.

VII. IDS IN VARIOUS DOMAINS

An IDS is used in numerous fields and the performance in each field is described and defines how they performed.

IDS in MANET Manetis defined as mobile adhoc network. It is an autonomous network that is composed naturally by the combinations of mobile nodes without centralized administration.IDS is used in Manet. Mobile network is normally needed in the battlefield for military people to get proper network [20]. Normally the messages are splited into number of packet and they use a hardware device like wire and modem to transmit. But, in Manet they are connected wirelessly. Watchdog and path rater are the two techniques added on the protocol in Adhoc. A watchdog identifies the misbehaving nodes by eavesdropping on the transmission of the next hop [20]. A path rater then helps to find the routes that do not contain those nodes [20]. IDS are used in Manet while transferring the series of packets to the destination through mobile network to find the intruder if any.

IDS FOR CLOUD COMPUTING Cloud computing is illustrated as internet based computing cloud where, virtual shared servers provide software infrastructure platform devices and other resources and hosting to customer as a service on pay-as you-use basis [21]. The user of the cloud does not hold any physical framework instead they lease from the mediator (third party). They pay only for the usage of the resource. Intrusion detection system plays an important role in the security and perseverance of active defense system against intruder hostile attacks for any business and IT organization [24]. In cloud computing the applications are received on the remote server of the provider and they have the control towards the usage of the data. IDMEF (Intrusion detection message exchange format) is the standard used in cloud for the communication purpose [21]. Cloud computing security issues Cloud data confidentiality Attacks on remote server Cloud security auditing Lack of data interoperability

IDS IN DATA MINING Data mining is the process of extracting the hidden knowledge from the databases. IDS are very important in data mining. Intrusion detection includes identifying a set of malicious actions that compromise the integrity and availability of information resources [22]. Intrusion detection in data mining has two divisions, they are, misuse detection and anomaly detection. In misuse detection the labeled data are built using anticipating model [23]. In anomaly detection there is a deviation between models. To use the data first it should be converted into International Journal of Security, Privacy and Trust Management (IJSPTM) Vol 4, No 1, February 2015 43 featured data and the data mining models are applied to it and they are summarized to produce the result.

VIII. TECHNICAL CHALLENGES

Large data size
Higher dimensionality
Data preprocessing

IX. CONCLUSION

The main objective of this paper is to provide an overview of the necessity and utility of intrusion detection system. This paper gives complete study about types of IDS, life cycle, various domains, types of attacks and tools. IDS are becoming essential for day today security in corporate world and for network users. IPS defines about the preventing measures for the security. In the lifecycle the phases developed and the stages are illustrated. Still, there are more challenges to overcome. The techniques of anomaly detection and misuse detection are specifically illustrated and more techniques can be used. Further Work will be done on comparative analysis of some popular data mining algorithms applied to IDS and enhancing a classification based IDS using selective feedback methods.

X. FUTURE SCOPE IN THE FUTURE

We recommend considering the Hybrid Intrusion Detection System which is better at detecting R2L and U2R attacks. The misuse detection approach better at detecting R2L and U2R attacks more efficiently as well as anomaly detection approach. Work for approach which is better at detecting attacks at the absence of match signatures as provided in the misuse rule files. The critical nature of the task of detecting intrusions in networks and applications leaves no margin for errors. The effective cost of a successful intrusion overshadows the cost of developing intrusion detection systems and hence, it becomes critical to identify the best possible approach for developing better intrusion detection systems. Every network and application is custom designed and it becomes extremely difficult to develop a single solution which can work for every network and application. In this thesis, we proposed novel frameworks and developed methods which perform better. However, in order to improve the overall performance of our system we used the domain knowledge for selecting better features for training our

models. This is justified because of the critical nature of the task of intrusion detection. Using domain knowledge to develop better systems is not a significant disadvantage; however, developing completely automatic systems presents an interesting direction for future research. The field of intrusion detection has been around since 1980's and a lot of advancement has been made in the same. However, to keep pace with the rapid and ever changing networks and applications, the research in intrusion detection must synchronize with the present networks. Present networks increasingly support wireless technologies, removable and mobile devices. Intrusion detection systems must integrate with such networks and devices and provide support for advances in a comprehensible manner.