IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITP5154

This paper appears in *Managing Modern Organizations Through Information Technology*, Proceedings of the 2005 Information Resources Management Association International Conference, edited by Mehdi Khosrow-Pour.

Use of Reconnaissance Patterns for Intelligent Monitoring Model

Mariana Hentea

Dept of Computer Sci. & Infor. Systems, Southwestern Oklahoma State University, Weathorford, OK, USA, mariana.hentea@swosu.edu

ABSTRACT

The increasing number of computer security attacks and intrusions affect organizations around the globe. This situation made security event management become mainstream important. One method called reconnaissance is used by hackers to choose networks and domains to search for targets before an attack. Reconnaissance allows a hacker to identify targets to be attacked or used for launching attacks. The targets are systems or networks with vulnerabilities. In order to protect against potential attackers, it is necessary to understand their reconnaissance methods and reasons. For example, by knowing the hacker's reconnaissance targets, network administrators and security staff can verify the targets and improve the security of the targets or the network. However, monitoring and analysis of hacker's reconnaissance patterns has to be done correctly and continuously to determine the impact they may have on the security management. Network administrators need automated and effective techniques for recognizing and analysis of the reconnaissance patterns.

The paper discusses a method for intelligent monitoring of the reconnaissance patterns, identification, and selection of hacker's reconnaissance patterns to be used as inputs to the security event management model.

SECURITY MANAGEMENT

Unauthorized accesses to personal records or organization's sensitive information are common cyber attacks. Cyber security plan launched by White House calls for more specific requirements for computer and network security as well as emphasis on the availability of commercial automated auditing and reporting mechanisms and promotion of products for security assessments and threat management (Rabinovitch, 2003), (Hwang, Tzeng, Tzai, 2003), (Leighton, 2004). The objective of security event management is the identification of attempted or ongoing attacks on a computer system or network. However, security event management is a continuous process and it is based on interdisciplinary techniques (Mena, 2003), (Maiwald, 2004). Advanced approaches based on Artificial Intelligence (AI) techniques are used to develop intelligent models for security event management.

INTELLIGENT MODEL

The development of intelligent models for security event management is focused on the selection and identifying of data needed to support useful feedback to network administrator or security staff. The objective of the model design is to concentrate on developing learning techniques that can support the business and advise the user to make decisions before the attack damaged the information or made the systems unavailable. Another important factor to consider is that systems, software, and security policies change themselves over time and across the different platforms and businesses. These issues require enhanced capabilities for the model such as to be adaptive. The model should be capable to support monitoring and control of the network to include data collected by all security technologies and network management systems instead of relying on data provided by a single system. Also, the model should be cost effective such that organizations could afford the use of advanced technologies for security protection (Geer, 2003).

An intelligent model based on AI techniques for cyber attack detection and prevention is discussed in (Hentea, 2004), (Hentea, 2003). Artificial Intelligence techniques such as data mining, artificial neural networks, fuzzy logic, and expert systems can be integrated with traditional procedural and statistical methods to analyze the collected data by sensors, recognize reconnaissance patterns, filter and correlate events to support security event management and prevention of intrusions. Data collected by sensors contain information about "certain activities that are often the precursor of an attack in progress" (Howlett, p. 235).

There are various classes of attackers and various types of attacks. Some classification schemes identify the attribute which is the basis of the classification such as the system component that was attacked, the intent of the attacker, the technique used in the attack, the reason why the exploited flaw is present in the system, the outcome of the intrusion, and others. Lindqvist and Jonsson (1997) propose a classification scheme based on taxonomy of intrusions; taxonomy of intrusions allows deriving statistics, observing patterns and drawing conclusions. Attack patterns are used in the attack modeling process (Cheung, Lindqvist, Fong, 2003). Also, there are projections of future sources of attacks using time series modeling method (Wei, 2004).

One method called reconnaissance is used by hackers to choose networks and domains to search for targets before an attack. Reconnaissance allows a hacker to identify targets to be attacked or used for launching attacks. The targets are systems or networks with vulnerabilities. In order to protect against potential attackers, it is necessary to understand their reconnaissance methods and reasons. For example, by knowing the hacker's reconnaissance targets, network administrators and security staff can verify the targets for vulnerabilities and improve the security of the targets or the network. However, monitoring and analysis of hacker's reconnaissance patterns has to be done correctly and continuously to determine the impact they may have on the security management. Reconnaissance patterns can be used to increase the detection and prevention capabilities of the security event management model. Next sections discuss reconnaissance patterns of attempts detected with intrusion detection systems and methods for intelligent monitoring of the reconnaissance probes including identification and selection of hacker's reconnaissance patterns to be used as inputs to the model.

RECONNAISSANCE ANALYSIS ISSUES

Intrusion detection systems, firewalls, anti-virus software, virtual private networks, encryption, and biometrics are security technologies that are common systems in use today. These systems generate hundreds of events and report various problems or symptoms. These systems are not efficient and scalable because they rely on human expertise to analyze periodically the data collected with all these technologies. The Intrusion Detection Systems (IDS) are software systems used to detect security intrusions to a network based on a number of signs called attack signatures. An IDS monitors the activity within a network and generates alerts if suspicious activity is detected. Also, an IDS can detect several reconnaissance activities. Reconnaissance phase involves collecting data about the target network or systems. Reconnaissance activities include scanning of ports, DNS queries and zone transfers, and browsing through the target's Web site. The efficiency of IDS depends on its

capabilities to detect reconnaissance activities and on the network administrator and security staff dedicating time and resources for analyzing collected data and understanding the meaning of the information. Volonino and Robinson (2004) suggest that "Careful inspection of the frequency, type, and source of attacks can lead to insights that the intrusion detection software cannot provide". Reconnaissance probes or recons can also be intermittent (called slow sweeps) to avoid detection. With data collected from the outside sensor, "it is possible to identify patterns showing new scans and attacks that are not captured by the IDS signature library" (O'Neil, 2003). A network administrator needs to know the techniques employed by attackers to stealth or obfuscate a scan and "basic methods to detect and defend against these recon attempts" (Millican, 2003) and perform "trend analysis of the alerts to determine what attacks are attempted and what machines/ports are scanned most often" (SANS Report, 2004). One defense method is blocking port scans. A network administrator will be able to increase the security level of the network. An effective method of blocking is to have a device or service at the network perimeter that recognizes scan packets and drops them. In addition, the data collected by an intrusion detection system is huge and significantly complex. Deriving meaningful information requires continuous analysis and correlation of events by network administrators. Therefore, these limitations combined with the attacks growth impact the efficiency of security management and increase the activities to be performed by network administrators. Organizations need a systematic and automated approach for security management that addresses security consistently at every level. Specific data processing issues include data collection, data reduction, data normalization, event correlation, behavior classification, reporting, and response (Hentea, 2005). Organizations need to implement an efficient reporting strategy that reduces the total cost of ownership (TCO) for an intrusion detection system. The following section describes scenarios and programs designed to analyze reconnaissance activities detected with SNORT, an open source for the intrusion detection system (SNORT, 2004), (Rehman, 2003).

SNORT Detection of Reconnaissance Activities

During a time period of less than a month, data was collected by three sensors based on SNORT software installed in the network of Purdue University Calumet. One SNORT sensor was installed on the external segment looking for any probes, scans, or attacks that are coming from the Internet. The other two SNORT sensors were deployed on the internal segments (faculty and students subnetworks). The collected data amounts to thousand of alerts classified in unique alerts and specific categories that were stored in a database requiring huge disk space. Although the software provides valuable reports regarding all alerts and recent alerts, categories, protocols, and detection time, it becomes difficult for a human to trace all source and destination (target) IP addresses reported. The following section is a summary of the manual activities involved in the analysis of the collected data related to reconnaissance activities.

Analysis of Reconnaissance Data

We identified a list of 25 IP distinct addresses that were frequently reported by SNORT as a target. Also, we determined the list of most active sources. The alerts indicated a number of persistent scans and attempts for unauthorized network access. In addition, the team identified that several attempts were from the same source. The attempts occurred periodically with a frequency varying from a high range to a low range or vice versa during a period of time spanning from hours to days. Although tracing the source is important, this activity was not performed because of the uncertainty of source IP addresses, since attackers can hide their identity by forging the source IP address.

The security staff had to translate the IP addresses using other sources of information in order to identify the targets or sources of network reconnaissance attacks. A team of three security staff spent several hours to monitor and trace a target IP address of a teardrop attack (a tool used for Denial of Service attacks). The team tried to understand the reason why the hacker has chosen a specific target. Identifying the target location and owner consumed significant time and resources.

Regardless of the hacker's motives, if it was a random target or if he was interested in a successful attack, we investigated and analyzed data captured by all security systems. The team verified the logs provided by other security technologies including the firewall. After learning the profile of the target – owner, type of computer, location, system software, and applications used, we concluded that patterns of attempts can help on understanding the hacker's motives for attacks. We investigated the security policies, executed tests, and monitored security events continuously. By monitoring the targets, we learned that we need to enhance the security measures for the targets. We identified requirements for tools to discover, recognize, and classify reconnaissance patterns used by hackers. Then, we correlated reconnaissance patterns with events and data generated by different security systems. It was observed that security systems lack the tools for the recognition, filtering, and classification of the reconnaissance patterns. Network administrators use a diversified set of software tools for monitoring the networks. Examples of such tools include network management systems, enterprise inventory of computers, network configurations. The data had to be extracted from these systems and then correlated to get useful information about the location and owner of the target, types of applications used, login and logout times, etc. This information was used to check the compliance of the targets with the security policies and to verify if the security policies were implemented correctly. In other words, the reconnaissance patterns helped on analyzing the efficiency of the security policies implementation as well as on reassessing the risks. In addition, the reconnaissance patterns helped on identifying the possible vulnerabilities that the hacker was trying to discover for later use. These reconnaissance patterns were then identified as inputs to an intelligent model. However, the isolated use of AI techniques for security event detection without integration with other techniques based on statistical and procedural methods provides limited support to a network administrator when decisions have to be made and quick actions to be taken to prevent or stop ongoing attacks.

Therefore, manual processing is time consuming, generates results after long delays, and is prone to errors. Consequently, automated processing tools are the only viable solution. Next section describes the methods used to automatically process reconnaissance data.

Processing Reconnaissance Data

The data needed for processing was extracted from logs, reports, and MySQL database using scripts. Then, the data was processed with programs written using MATLAB software. The programs support the trend analysis and plotting of graphs such as the frequency, type, targets, and source of attacks. In addition, programs were used to select and classify the targets that were attacked continuously by sources using the same tool or different tools. The reports provide comprehensive information about the targets and sources that the intrusion detection software did not generate. This information allowed identifying reconnaissance aspects such as targets with the highest number of scans, frequency, and period of time the reconnaissance activities occurred. It was determined that the patterns of reconnaissance events are described by a set of characteristics such as frequency, number of sources attacking the same target, duration, target profile, source, type of cracking tools, and number of different cracking tools used for a target, source attacks against other targets. These characteristics provide useful information to a network administrator to verify the security countermeasures and reassess the vulnerabilities for a monitored network.

Figures 1-6 show attempts for intrusion in the network and computers that were identified for the network during a period of time. It is observed that a source could be very active by launching several attacks by using different tools against one or more targets during the same period of time. The charts depicted in Figures 1-3 are based on data collected during a period of five days and charts depicted in Figures 4-6 are plotted based on data collected during a period of two weeks. Figure 1 shows the attacks distribution from the same source and Figure 2 shows the attacks

occurrence per day from the same source. Figure 3 shows the attacks distribution per target. It is concluded that a hacker was attempting to penetrate the organization and within organization he was trying to damage specific targets. Profiles of attack patterns are being used as inputs to the intelligent model for security event management.

To protect against potential attacks, it is essential that security event management model includes reconnaissance patterns to aid in the process of detecting attacks and predicting new attacks. These reconnaissance patterns are included as inputs to the module based on artificial neural networks. An expert system module is used to classify the tools available to hackers and build knowledge regarding security policies and operational security for the targets. The tools used by hackers can facilitate the process of securing the network. This information allows identifying defensive measures that will reduce hacker's reconnaissance effectiveness. Documentation about evolving threats is important in making timely adjustments to security policies Such that the basic security goals are achieved in an efficient manner (Stoneburner, Hayden, Feringa, 2001).

Both artificial neural network module and expert system module are components of the intelligent model. The following section discusses the advantages of advanced Artificial Intelligence techniques for the security intrusion prevention.

SIGNIFICANCE OF AI TECHNIQUES

Technologies such as network intrusion detection systems used to detect intrusion attempts lack the capability to analyze and generate useful information to a human to take prompt actions and prevent the intrusion. There should be automated tools that monitor reconnaissance activities and notify the human if the targets are in danger of being attacked. AI techniques are being used in intelligent models to improve the efficiency of security event management. AI techniques such as data mining, artificial neural networks, fuzzy logic, and expert systems can be integrated with traditional procedural and statistical methods to analyze the collected data by sensors, recognize reconnaissance patterns, filter and correlate events to support security event management and prevention of intrusions. These techniques improve the ability of security systems to correlate events generated by a diversified suite of modern tools used for network management and security monitoring.

Although the intelligent model is generic, specific solutions based on AI approaches can be derived to support the business needs of each enterprise. One technique, called feature selection, is used to reduce overhead and improve classification of events by decreasing the amount of information required to make good classifications and predictions. Data mining is often defined as finding hidden information in a large amount of data. However, the Artificial Intelligence techniques involve areas outside data mining.

Machine learning technique is concerned with writing programs that can learn and adapt in real time. This means that the computer makes a prediction and then, based on the feedback as to whether it is correct, learns from this feedback. It learns through examples, domain knowledge, and feedback. When a similar situation arises in the future, this feedback is used to make the same prediction or to make a completely different prediction. The results of the predictions must be significant and must perform better than statistical predictions. The effects on the user and security management efficiency are analyzed using modeling techniques.

Figure 1. Number of Attacks by Source

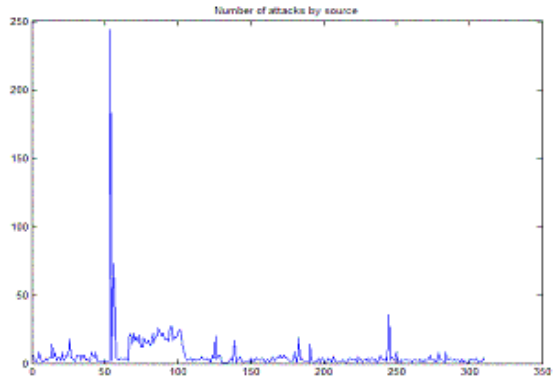


Figure 2. Number of Attacks/Day from Most Active Source

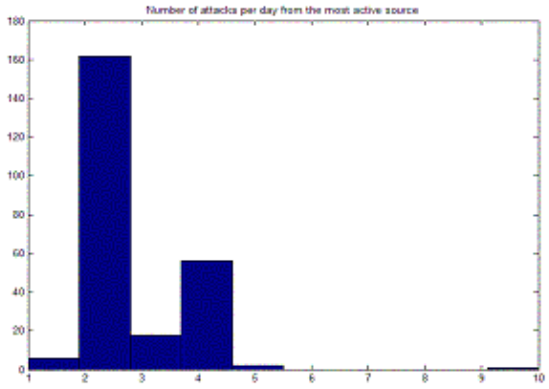
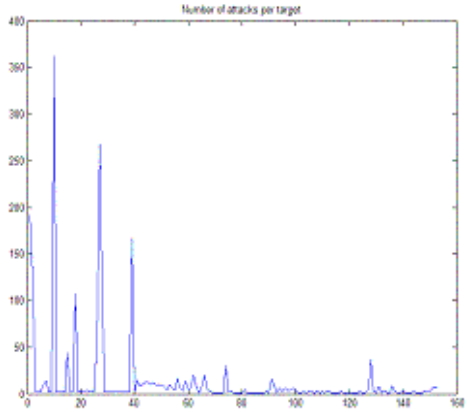


Figure 3. Number of Attacks per Target



CONCLUSIONS

In this paper we discussed a framework for intrusion mining as the application of data mining techniques to discover reconnaissance patterns from attack attempts. It is not sufficient to know the types of intrusions as provided by intrusion detection systems. There is a need to discover the reconnaissance patterns to enhance the security of the system and provide useful information to the network administrator.

By analyzing the logs of traffic, usually on the outside interface, it is possible to identify patterns showing new scans and attacks that are not captured by the IDS signature library. In addition, these analyses make possible to answer many questions about the effectiveness of the firewall, and the kinds and volume of traffic flowing through the network.

Future work will seek to integrate the model with the security policies and define the mechanisms for communication and sharing the informa-

Figure 4. Attacks from the Second Active Source

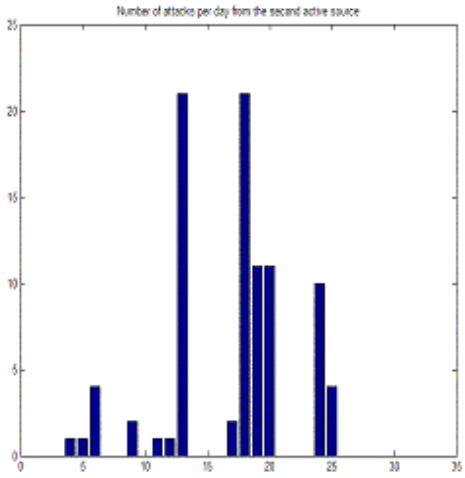


Figure 5. Attacks from the Most Active Source

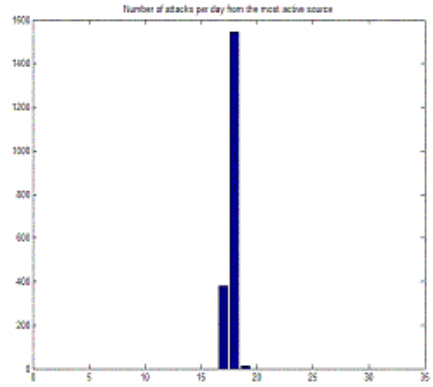
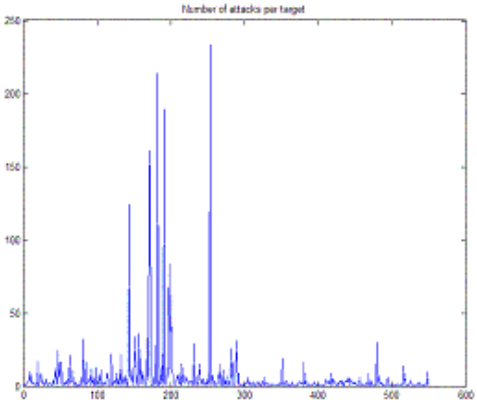


Figure 6. Distribution of Attacks per Target



tion with the humans. The model should include input from the security plans and measures for network monitoring, auditing, physical and logical access control. The model can be used for developing tools based on agent technology to monitor, detect, and identify the threats as well as prevent attacks by providing useful information to a network administrator before the attack occurred.

REFERENCES

Cheung, S., Lindqvist, U., Fong, M.W. (2003). Modeling Multistep Cyber Attacks for Scenario Recognition. Proceedings of the Third DARPA Information Survivability Conference and Exposition (DISCEX III), Washington, D.C. (1), 284-292.

Geer, D. (2003). Risk Management Is Still Where the Money Is. IEEE Computer, 36(12), 129-131.

Hentea, M. (2003). Intelligent Model for Cyber Attack Detection and Prevention. Proceedings of the ISCA 12th International Conference Intelligent and Adaptive Systems and Software Engineering, July 2003, San Francisco, California. 5-10.

Hentea, M. (2004). Data Mining Descriptive Model for Intrusion Detection Systems. Proceedings of 15 Information Resources Management Association International Conference, New Orleans, Louisiana, May 2004.

Hentea, M. (2005). Information Security Management. Chapter in Encyclopedia of Multimedia Technology and Networking. Editor M. Pagani. IDEA GROUP Inc.

Howlett, T. (2004). OPEN SOURCE SECURITY TOOLS. A Practical Guide to Security Applications. Prentice Hall, Upper Saddle River, New Jersey.

Hwang, M-S. Tzeng, S-F., Tsai, C-S. (2003). A New Secure Generalization of Threshold Signature Scheme. Proceedings of IEEE Conference on Information Technology for Research and Education, August 2003, Newark, New Jersey. 282-285.

Leighton, F.T. (2004). Hearing on the State of Cyber Security in the United States Government. Computer Security Journal. XX (1), 15-22.

Lindqvist, U. and Jonsson, E. (1997). How to Systematically Classify Computer Security Intrusions. Proceedings of the 1997 IEEE Symposium on Security & Privacy, Oakland, California, May 1997. 154-163.

Maiwald, E. (2004). Fundamentals of Network Security. McGraw Hill, New York, New York.

Mena, J. (2003). Investigative Data Mining for Security and Criminal Detection, Butterworth Heinemann, Amsterdam, Netherlands.

Millican, A. (2003). Network Reconnaissance – Detection and Prevention, GSEC v1.4b, SANS Institute 2003. http://www.giac.org/practical/GSEC/Andy_Millican_GSEC.pdf. Accessed 12/05/04.

O'Neil, P. (2003). Build your own firewall using SuSE Linux: A Mechanics guide, version 2.5b. SANS Institute 2003. www.giac.org/practical/GSEC/Paul_ONeil_GSEC.pdf. Accessed 12/05/04.

Rabinovitch, E. (2003). Maintaining a Secure Networking Infrastructure. Proceedings of IEEE Conference on Information Technology for Research and Education, August 2003, Newark, New Jersey. 587-589.

Rehman, R.U., INTRUSION DETECTION with SNORT, Prentice Hall, 2003, Upper Saddle River, New Jersey.

SANS Report, http://frasier.dpo.uab.edu/security/sansreport. Accessed 10/3/2004.

SNORT. Release 2.1.3. http://www.snort.org

Stoneburner, G., Hayden, C., Feringa, A. (2001). Engineering Principles for Information Technology Security (A Baseline for Achieving Security). NIST Special Publication 800-27, June 2001.

Volonino, L., Robimnson, S.R. (2004). Principles and Practice of INFORMATION SECURITY. Pearson Prentice Hall, Upper Saddle River, New Jersey.

Wei, J. (2004). Forecasting on Information Attack Sources. Proceedings of 15 Information Resources Management Association International Conference, New Orleans, Louisiana, May 2004. 961-963.