Attacker Behavior Analysis in Multi-stage Attack Detection System

Rajeshwar Katipally
University of Tennessee at Chattanooga
615 McCallie Avenue
Chattanooga, TN 37403
1 423 987 4629
Rajeshwar-katipally@utc.edu

Li Yang
University of Tennessee at Chattanooga
615 McCallie Avenue
Chattanooga, TN 37403
1 423 425 4392
Li-Yang@utc.edu

Anyi Liu
Department of Computer Science
George Mason University
Fairfax, VA 22030
aliu1@gmu.edu

ABSTRACT

Today's internet world is facing attacks from different types of attackers who launch is multistage attack. Besides discovering, visualizing, and predicting multi-stage attacks, a method to understand and profile behaviors of attackers is important to protect network security. We use the Hidden Markov Model (HMM) to analyze and predict the attacker behavior based on what was learned from observed alerts and intrusions. We use data mining to process alerts to generate input for the HMM to calculate the required probability distribution. Our system is able to stream real-time Snort alerts and predict intrusions based on our learned rules. Our system is able to automatically discover patterns in multistage attack, classify attackers based on their behavior pattern. By doing this, our system can effectively predict behavior and attackers and assess danger level of different groups of attackers.

Categories and Subject Descriptors

C.2.3 **[Computer Communication Networks]:** Network Operations - *Network Monitoring*; K.4.4 **[Computers and Society]:** Social Issues - *Abuse and crime involving computers;* K.6.5 **[Management of Computing and Information Systems]:** Security and Protection – *Unauthorized Access.*

General Terms

Security

Keywords

Intrusion Detection, Behavior Analysis, Multi-stage Attack.

1. INTRODUCTION

Multi stage attacks are achieved gradually in multiple steps. Attackers gain knowledge of the target system in each stage, which prepares for the next stage attack. The attacker tries to collect information and exploit vulnerabilities of the target system in initial stages, and penetrate into the target system to compromise the resources of the targeted system in the latter stages. For example an attacker attempts to gain access to an organization's database. First, he has to find the entry points to that organization, which can be known by scanning. Then he will have to find how he can access the database such as password cracking or SQL injection. It is important to discover behavior pattern of attackers to design an effective IDS to detect multi-stage attacks.

Reports from U.S. Government Accountability Office [1] have classified attackers in to different set of groups such as criminals, terrorists, and insiders. Criminals attempt intrusion for money so that they try to hide their identity. Most dangerous kind of attackers is insiders. Insiders usually might not have knowledge about intrusions, but they know everything about the organization to gain unrestricted access to the resources. Insiders can be contractors hired by organization or employees who accidentally insert malware into the system [1]. Terrorists are the attackers with intention of destroying the infrastructure and killing the citizens.

Predicting the type of an attacker based on his/her behavior is important to reduce risks and damage of computer systems. However, there is no effective model or tool that can classify attackers based on their behaviors in multi-stage attacks. We need develop a novel system that is able to classify type of attackers and therefore predict their behaviors. To achieve this goal we have process alerts from Snort, a rule based IDS to profile behaviors of different attackers. These profiles are used to detect the multistage attacks, whereas to analyze the attackers' behaviors and intentions using the Hidden Markov Model (HMM). The HMM generalizes the steps in a multistage attack to analyze the ongoing behaviors of attackers. To implement the behavior analysis we have also used Kullback Leibler Distance Calculator [9], which determines the most similar one with the five stored behavior models. Our system are able to efficiently remove redundant alerts, discover temporal patterns of attacker behaviors, and classify attackers based on their behaviors and predict next stage of an attack.

2. RELATED WORK

Multistage attack detection is a relatively new and very challenging area in the domain of network security. Valuable contribution made by some domain experts such as Mathew et al [3], Kumar [2], and Wanderer [8]. Mathew et al [4] have made a good effort to present a technique for understanding multi stage attacks using attack-track based visualization of heterogeneous event streams. They have developed a technique that provides the security analyst with real-time multistage attack visualization. They have used IDS alerts and system log files as the basis for making assumptions about multi stage attacks and they have classified these sensor events based on semantic contents and alert classification scheme presented in [3]. They have used event correlation [4] to find the multistage attack, in order to prevent event correlation from considering every occurred event to find the relation between the alerts, which is sometimes not effective due to false detection by IDS; they have proposed a guided event fusion technique. IDS sensors usually outputs events based on the detection technique they follow. For sensors that perform misuse detection the event set contains the events that are matched with signature set that sensor uses to perform intrusion detection, whereas for anomaly detection the set contains all the events that deviates from normal activity. For both techniques the IDS sensors generate diverse and finite set of events. Multilevel alert clustering in [7] and intelligent alert clustering model in [6] are well formed techniques for eliminating the false alarms. For analyzing the attackers' behavior it is very important to classify the incoming alerts into groups that most effectively indicate the stage in a multistage attack. In [3] the alert classification scheme handles this problem up to some extent. The event fusion technique proposed in [3] basically tries to correlate the events together and combines them into certain number of multi stage categories such as Reconnaissance, Intrusion, Privilege Escalation and Goal. Each of these stages has its own weight to indicate how important that stage is in multistage attack. As these four categories are not enough to analyze the behavior we have increased the number of categories to five. They are Scanning, Enumeration, Access attempt, Malware attempt, and Exploitation- denial of service. Oursron et al. used the Hidden Markov Model (HMM) for detecting multistage attacks [5]. They have defined some rules based on the available data. They have defined each state of the HMM as a stage in a multistage attack. Transitions among the states are governed by a set of probabilities with each state. Wang et. al. [5] proposed a real-time unified approach to correlating, hypothesizing, and predicting multi-step attacks based on networking intrusion alerts. Their approach was based on the assumption that a graph, which describes vulnerabilities and connectivity of a network, can be generated from a vulnerability scan. In addition, they cannot categorize the motivation of attackers from the observed events. Finally, they cannot discovery attacking patterns automatically.

3. MULTI-STAGE ANALYSIS

Different group of attackers have different motivation to perform an attack which results in various level of damage. Attackers combine different stages sequentially to launch intrusions. The outcome of one stage serves as an input to its subsequent stage. We analyzed and defined the stages in a multistage attack as reconnaissance, scanning, enumeration, exploitation, access attempt, malware attempt, and exploitation denial of service.

Reconnaissance and Scanning is the stage by which a potential attacker or intruder learns all the information needed about the network or system to use it in further stages. It is an unauthorized discovery of vulnerabilities of the targeted network. The intruder tries to send more effective request to the target system to gather more useful information. In this stage intruder tries to know which ports are open, which resources are accessible, and what the vulnerable points are in the targeted system. The scanning can be done by either network mapping or port scan or vulnerability scan. The common network mapping attacks are ICMP trace route, ICMP echo request, ICMP ping, ICMP fingerprinting and TCP/UDP scan.

Enumeration: After scanning there is another stage called enumeration. In enumeration intruder continuously keeps on scanning the targeted system to exploit the vulnerabilities. An example of enumeration is DNS request, where intruder sends repeated requests to the domain name server.

Exploitation by Access Attempt: After gathering the information about the entry points and vulnerabilities of the system, the next stage is access attempt. In this stage the intruder tries to gain access to the resources of the system or network to misuse them. Once the intruder finds the vulnerabilities in the system by scanning, he tries to gain the access control to the resources that are required for the latter stages of the attack. An access attempt can be done by an outsider or insider. Access control is required because without some special or administrator privileges, it is not possible to compromise the targeted network's resources. There are a number of attacks to gain access control. Buffer overflow, SQL injection and SQL private access are few of the many attacks. Buffer overflow is one of the well-known attacks to control the access of the target machine. Buffer overflow is something that an attacker tries to write more than a program can accommodate on the target machine. By doing so attacker gets the control of that particular application and from then on attacker can execute whatever he would like to. With SQL injection intruder tries to inject own SQL code to gain some privileges.

Exploitation by Deny-of-Service: A possible next step is exploitation-denial of service. An intruder can directly go for denial of service attack after the scanning attack, or he can also go through the access attempt before the denial of service attack. Denial of service attack is one of the most dangerous attacks in the world. There have been records of denial of service attacks against famous social networks such as Twitter and Facebook, and also some government agencies. From scanning stage an intruder know which ports are open and the capacity of target system or server. Based on that information intruder floods the requests to the server more than what it can handle, in that way intruder denies any other user requests to that particular server. One other way to do denial of service attack is by gaining the access privileges to the targeted system and makes the system inaccessible. There are many tools in the market for denial of service attack. Tools like Trino, The Tribe Flood Network are very effective for distributed denial of service attacks.

Exploitation by Malware Attempt: Another stage in the multi stage attack is Exploitation-malware attempt, where attacker tries to execute his own code on the target machine to make the resources compromise. This step is the most dangerous step in multi stage attack. In this stage attacker tries to achieve whatever he wants. By the time attacker enters this stage he'll have all the information about the target machine and he'll also have the access permissions to the resources that he wants to make compromise. There are some powerful attacks to achieve the purpose in this stage. Attacks such as Heap overflow, SQL injection are few of the many attacks of this type. An attacker after performing a buffer overflow attack, he gains the total control of the application on target machine then he tries to insert his own executable code in to that program. When the application runs normally it executes the attacker's code on the target machine, which eventually might cause a serious damage to the target machine

4. ANALYZING ATTACKER BEHAVIORS

4.1 Alert Analysis

We have repeated well-known intrusions to our target network by emulating different groups of attackers. The collected traffic was used to learn rules and profile behavior pattern of multistage attackers. We have defined five stages in our model: scanning, enumeration, exploitation by access attempt, exploitation by denial of service, and exploitation by malware attempt. Alerts are mapped into five different sets which are the same as states in our model as discussed in Section 3. For example an alert of *ICMP PING* type is usually considered as a scanning type and an alert of *SHELLCODE X86 INC EXC NOOP* is considered as exploitation malware attempt type. In this way we have defined rules to train our model. As of now we have around 65 rules to train our model. Once the rule set is defined, we have assigned the state name to each alert by applying rules. For example an alert like below

07/14-13:12:54.775367 [**] [1:384:5] ICMP PING [**] [Classification: Misc activity] [Priority: 3] {ICMP} 192.168.1.24 -> 192.168.1.1

is converted to

Scanning, 07/14-13:12:54.775367, Misc activity, 192.168.1.24, 192.168.1.1

The main purpose of converting the alerts into this form is they are sent to HMM to train our system for different type of attackers.

4.2 Training System Using Hidden Markov Model

Now we have all the data that we need to train our system. We have formed 5 set of alert sets that match 5 states of our model. By doing a lot of research we have defined the attacker behavior based on their intension and level of expertise. Based on their intentions they are divided into following 8 groups.

Criminal groups seek to attack systems for monetary gain. Specifically, organized criminal groups use spam, phishing, and spyware/malware to commit identity theft and online fraud. International corporate spies and criminal organizations also pose a threat to the United States through their ability to conduct industrial espionage and large-scale monetary theft and to hire or develop hacker talent.

The disgruntled organization **insider** is a principal source of computer crime. Insiders may not need a great deal of knowledge about computer intrusions because their knowledge of a target system often allows them to gain unrestricted access to cause damage to the system or to steal system data. The insider threat includes contractors hired by the organization, as well as employees who accidentally introduce malware into systems.

Terrorists seek to destroy, incapacitate, or exploit critical infrastructures in order to threaten national security, cause mass casualties, weaken the U.S. economy, and damage public morale and confidence. Terrorists may use phishing schemes or spyware/malware in order to generate funds or gather sensitive information.

Hackers break into networks for the thrill of the challenge, bragging rights in the hacker community, revenge, stalking others, and monetary gain, among other reasons. While gaining unauthorized access once required a fair amount of skill or computer knowledge, hackers can now download attack scripts and protocols from the Internet and launch them against victim sites.

Phishers are individuals, or small groups, who execute phishing schemes in an attempt to steal identities or information for monetary gain. Phishers may also use spam and spyware/malware to accomplish their objectives.

Nations use cyber tools as part of their information-gathering and espionage activities. In addition, several nations are aggressively working to develop information warfare doctrine, programs, and capabilities. Such capabilities enable a single entity to have a significant and serious impact by disrupting the supply, communications, and economic infrastructures that support military power—impacts that could affect the daily lives of U.S. citizens across the country.

Spyware/malware authors are individuals or organizations with malicious intent to carry out attacks against users by producing and distributing spyware and malware. Several destructive computer viruses and worms have harmed files and hard drives such as the Melissa Macro Virus, the CIH (Chernobyl) Virus, Nimda, Code Red, Slammer, and Blaster.

Bot-net operators use a network, or bot-net, of compromised, remotely controlled systems to coordinate attacks and to distribute phishing schemes, spam, and malware attacks. The services of these networks are sometimes made available on underground markets (e.g., purchasing a denial of service attack or servers to relay spam or phishing attacks).

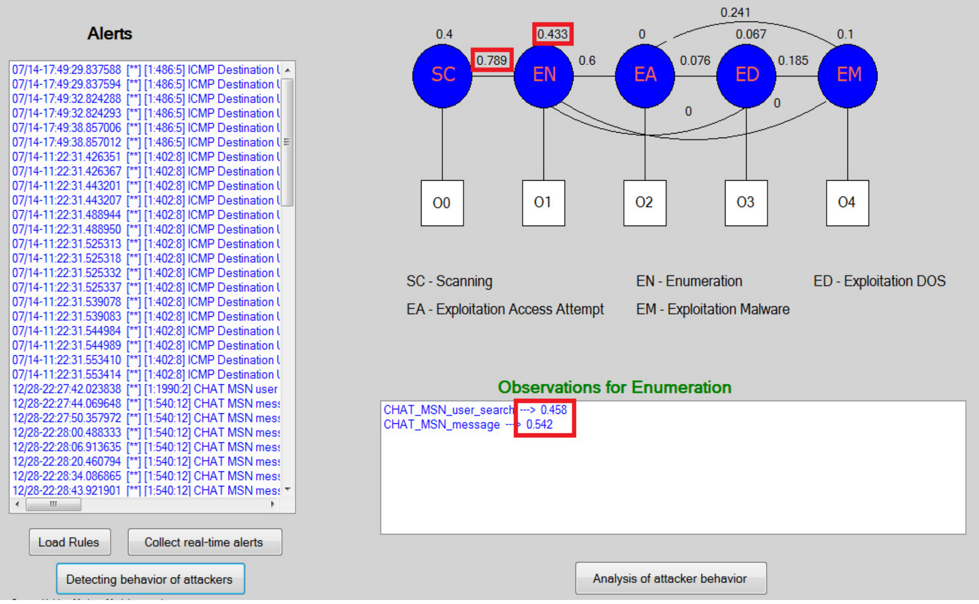
To calculate the probabilities for each type of the behavior using HMM we have created 5 sets of alert each one of them represents one of the above classified behaviors.

4.3 Hidden Markov Model

The HMM used to model motivation of attackers is defined as λ= (A, B, π, N), Where N is 5, state probabilities π is 1 × 5, transition probabilities A = {aij} is 5 × 5, and observation probabilities B = {bj(k)} is 5 ×M (M is the number of different attack types). After initializing the model, forward algorithm uses initial state and transition probabilities to calculate the observation probabilities for each state based on occurred observation sequence. The backward algorithm uses the calculated observation probabilities and processes backward to calculate state and transition probabilities for each state.

5. IMPLEMENTATION

To show attacker behavior analysis we designed a simple interface that displays the alert that we have loaded on to the system and a fixed model which looks like a 5 state HMM. We have three buttons one for each loading predefined rules, loading alerts, and to calculate the probabilities. To train the system, first, the predefined rules needs to be provided to the model to assign state to each of the alerts that would be loaded based on their alert type. Once the rules are loaded we have to load a set of alerts that are manually grouped together for each type of the 9 behaviors. The system takes each set of alerts and processes them through and generates a file that contains alerts with assigned state names (which are considered as states for HMM) and corresponding observations. Once this file is generated system takes it and calculates the probabilities using the HMM. Once the probabilities were calculated for each of the behavior types, we stored them to analyze the attacker behavior in the future.



```
07/14-17:49:29.837588  [**] [1:486:5] ICMP Destination U
07/14-17:49:29.837594  [**] [1:486:5] ICMP Destination U
07/14-17:49:32.824288  [**] [1:486:5] ICMP Destination U
07/14-17:49:32.824293  [**] [1:486:5] ICMP Destination U
07/14-17:49:38.857006  [**] [1:486:5] ICMP Destination U
07/14-17:49:38.857012  [**] [1:486:5] ICMP Destination U
07/14-11:22:31.426351  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.426367  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.443201  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.443207  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.488944  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.488950  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.525313  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.525318  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.525332  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.525337  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.539078  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.539083  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.544984  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.544989  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.553410  [**] [1:402:8] ICMP Destination U
07/14-11:22:31.553414  [**] [1:402:8] ICMP Destination U
12/28-22:27:42.023838  [**] [1:1990:2] CHAT MSN user
12/28-22:27:44.069648  [**] [1:540:12] CHAT MSN mess
12/28-22:27:50.357972  [**] [1:540:12] CHAT MSN mess
12/28-22:28:00.488333  [**] [1:540:12] CHAT MSN mess
12/28-22:28:06.913635  [**] [1:540:12] CHAT MSN mess
12/28-22:28:20.460794  [**] [1:540:12] CHAT MSN mess
12/28-22:28:34.086865  [**] [1:540:12] CHAT MSN mess
12/28-22:28:43.921901  [**] [1:540:12] CHAT MSN mess
```

SC - Scanning
EN - Enumeration
ED - Exploitation DOS
EA - Exploitation Access Attempt
EM - Exploitation Malware

```
CHAT_MSN_user_search ---> 0.458
CHAT_MSN_message ---> 0.542
```

Figure 2 Probability Distribution for an Ongoing Attack

Once the models were stored in the database, system was ready to use. A random alert set was given as an input our system. It calculated the probability distribution for the given set of alerts using the same approach we followed to train the system. As shown in Figure 2 generated HMM model has higher probabilities for scanning and enumeration, which indicated that the attacker was performing more initial stage operations. As indicated in observation textbox the attacker was performing CHAT_MSN_user_search and CHAT_MSN_message attacks. The calculated probability distribution was compared with each of the 5 stored probability distributions using the Kullback-Leibler distance calculator. In this example calculated probability distribution was closest to the model that was designed for terrorists, hackers, and criminals. Therefore the attacker could be one of the three types of the attackers. Below is the screenshot that shows a behavior type and its description.

6. Conclusion

We are able to profile and classify attackers in to different groups to measure serious levels of ongoing attacks, which provide valuable insight in how to protect computer and network systems.

7. Acknowledgement

This work is partially supported by NSF DUE-0942581, DUE-1129444, and Tennessee Higher Education Commission's Center of Excellence in Applied Computational Science and Engineering.

8. REFERENCES

[1] Cyberspace: United States Faces Challenges in Addressing Global Cybersecurity and Governance, GAO-10-606, URL: http://www.gao.gov/products/GAO-10-606 , July 2010.

[2] Dokas, Kumar, Lazarevic, Srivastava, Tan, Data Mining for Network Intrusion Detection, USA, 2004.

[3] S. Mathew, D. Britt, R. Giomundo, S. Upadhyaya, S. Sudit, Real-time Multistage Attack Awareness Through Enhanced Intrusion Alert Clustering, *In Situation Management Workshop (SIMA 2005), MILCOM 2005*, Atlantic City, NJ, October, 2005.

[4] Mathew, Giomundo, Uoadhyaya, Sudit,Slotz, Understanding multistage attacks by attack-track based visualization of heterogeneous event streams, *Proceedings of the 3rd international workshop on Visualization for computer security*, Virginia, USA, 2006.

[5] Lingyu Wang, Anyi Liu, Sushil Jajodia. Using attack graphs for correlating, hypothesizing, and predicting network intrusion alerts. *Computer Communications*, Vol.29, No.15, 2006, pages 2917- 2933.

[6] Ourston, et all, Applications of Hidden Markov Models to Detecting Multi-stage Network Attacks, *Proceedings of the 36th Hawaii International Conference on System Sciences*, 2003.

[7] Siraj, Maarof, Zaiton, Hashim, Intelligent Alert Clustering Model for Network Intrusion Analysis, *ICSRS Publication*, 2009.

[8] Siraj, Vaughn, Multilevel Alert Clustering for Intrusion Detection Sensor Data, Fuzzy Information Processing Society, USA, 2005.

[9] C.Warrender, S. Forrests, and B. Pearlmutter. Detecting intrusions using system call: Alternative data models. *In proceedings of the 1999 IEEE Symposium on Security and Privacy*, May 1999.

[10] S. Kullback, R. A. Leibler: On information and sufficiency, Annals of Mathematical Statistics, 22(1):79–86, 1951.