Cyber Threat Analysis for University Networks via Virtual Honeypots

Will Davis III
Florida Agricultural and Mechanical University
Department of Computer and Information Sciences,
Tallahassee, FL 32307-5100
will3davis@yahoo.com

Hongmei Chi
Florida Agricultural and Mechanical University
Department of Computer and Information Sciences,
Tallahassee, FL 32307-5100
hchi@cis.famu.edu

ABSTRACT
Multi-system university computer laboratories are becoming more vulnerable to attacks as the gain for the attackers become more lucrative. Universities are a special case of the traditional commercial networks as they are constantly under attack due to their large number of desktops and large data pipes to the Internet and other research institutions. There are numerous of malicious software created for the purpose of harming computer systems. With new attacks, it is difficult to defend networks against the attack until you determine how it was created. Therefore, within university networks it is urgent to identify immediately after one computer has been compromised.

In this project, I will examine cyber threats of university computer labs and investigate the features of various honeypots hosted on a virtual machine. I will implement multiple real-world Virtual Honeypots to compare their strengths and weaknesses to select the appropriate solution for implementation in a university network. These honeypot solutions are designed to collect data from attackers to determine if a computer within that network has been compromised.

Categories and Subject Descriptors
D.3.3 [**Information Security**]: Protecting information and the systems which hold that information from unauthorized access, use, disclosure, disruption, modification, perusal, inspection, recording or destruction.

General Terms
Cyber Analysis and Security

Keywords
Virtual, honeypot, benchmark, botnet, virtual machines, compromised machines, cyber threat analysis

1. INTRODUCTION
Information Assurance is currently one of the most trending topics in the field of computer science. It is defined as the managing risks related to the use, processing, storage, and transmission of information and data and the systems and processes used for those purposes. The five principles for protecting and defending information systems are to ensure *confidentiality* (nondisclosure of certain information except to an authorized person), *integrity* (dedicated adherence to a strict moral or ethical code), *authentication* (determining whether someone or something is, in fact, who or what it is declared to be), *availability* (the quality of being at hand when needed), and *non-repudiation* (guarantee that someone cannot deny something).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. 49th ACM Southeast Conference, March 24-26, 2011, Kennesaw, GA, USA. Copyright 2011 ACM 978-1-4503-0686-7/11/03.....$10.00.

The subject of intrusion detection systems was developed by James Anderson's paper, Computer Security Threat Monitoring and Surveillance in the early 1980's. [2]Anderson's paper introduced the notion that audit trails contained vital information that could be valuable in tracking misuse and understanding user behavior. The contributions of Anderson provided the earliest audit systems which enabled companies to monitor the activities of their network both inside and outside the company's infrastructure. The intrusion detection system is placed in-between an unmonitored internet connection and firewall. This gives the system the ability to capture the maximum amount of data since the packets are not being filtered. As a secondary precaution, another intrusion detection system can be placed "behind" the firewall to catch data that has been missed or initially deemed safe.

2. VIRTUAL MACHINES
When a virtual environment is loaded directly on top of an existing operating system, it is referred to as an Application Based VM. This type of architecture is also known as a Hosted VM. Hosted VMs are installed on a computer just as any other application program [3]. An advantage with this approach is the virtual environment may use the resources of the operating system (i.e. Device Drivers) for functionality. One disadvantage associated with Hosted VM's is the processor of the hardware has to keep up with both the real and virtual environments which limits the VM's efficiency. This problem is intensified when deciding to run multiple virtual environments. For this reason, Hosted VMs are typically only chosen when a single virtual environment is needed.

System Based Platforms are characterized as VMs that are loaded directly on the hardware as its own operating system. This is typically always the best approach due to the elimination of the host operating system which enhances the VM's extensibility and efficiency. The resources of the hardware can directly be focused on the VM and each guest operating system can share services equally [3]. However, a disadvantage to this approach is the resources (i.e. Device Drivers) will have to be separately installed on the VM as well as each one of VM's guest operating systems.

It may be difficult to find these resources for newer technology. Figures 2 and 3 are representations of the layered structures of Hosted and System VMs.

For the purposes of this project, the VM will need to run multiple guest operating systems. This will require the VM to extremely efficient. This project will also need a VM that will allow future guest operating systems to be added to it for future work. Considering the capabilities of both implementation techniques, the most suitable type of implementation for this project will be a System Based VM.

After examining the specification of different virtual machines based on their characteristics, the most suitable VM for this project is VMware. The specific server application of VMware utilized will be VMware Server ESXI. The biggest factors in determining which server application to use came down to guest operating systems, speed, supported drivers and USB support. These are the key factors that will be used constantly while constructing and implementing this project.

3. BOTNET ATTACKS

For this project, the cyber threat analysis will be focused around the practices of Botnet attacks. The evolution of bots stemmed from its original use in the Internet Relay Chat (IRC) which is a system that provides instant messaging over the Internet. These bots were later used in a malicious way to perform Denial of Service (DoS) attacks. Also known as a zombie army, a Botnet is defined as a host of internet enabled computers who have been compromised and configured in a way that will allow the set up of forwarding instructions to other internet enabled computers [1]. These compromised computers are referred to as zombies and are configured in such a way that prevents the host from knowing it is infected. Bonets were initially used maliciously because the leader could hide itself since the DoS attacks came from the infected machines. The zombie's roll in the attack is to serve the wishes of some master spam or virus originator [1]. The originator's objective is to collect as many zombies as possible to strengthen its army.

4. UNIVERSITY NETWORK ANALYSIS

Enterprise networks such as university computer labs are normally composed large numbers of workstations (computers) to meet the needs of large groups of students and faculty. These workstations are also located in multiple areas around the university. For leaders of botnet attacks, this is a perfect solution to gain multiple resources needed for creating a supercomputer aimed at performing a malicious attack. Because these large computer labs may have hundreds of workstations in them, the attacker can afford to take minimal resources from each workstations which makes it difficult for system administrators to notice a difference in computer performance. This in turn makes it difficult for the system administrators to defend the labs against attacks.

Another issue associated with university networks are student and personnel turnover. Stakeholders of these networks are constantly entering and leaving the campus and with each new stakeholder, a new security risk is presented. This is due to the fact that stakeholder has different needs which can lead to them lowing the security standards of computers located on the network. Also many of the personnel (university instructors) utilize resources across open networks for their research and course tools. This can lead to open ports and disabled firewalls for which the attacker is provided access into the network..

5. A SET OF HONEYPOTS

There will be three unique low interaction honeypots developed. Along with the honeypot architecture, each honeypot will be running separate data collection software tools with its own hostname and IP address. By including additional network defense software to the Honeypot solutions, it will become possible to gain more accurate data as well as gain a perspective of malicious data like never before. Table 1 is a look at some of the characteristics of the honeypots that will be developed for this project.

Table 1: Honeypot Architectures

#	Guest O.S.	Honeypot Solution	Data Collection/Analysis Tools
1	Windows XP SP2	Honey@home	CWSandbox, Tiny Honeypot, TCP Dump
2	Windows XP SP2	HoneyC	CWSandbox, Tiny Honeypot, TCP Dump
3	Windows XP SP2	Honeysnap	CWSandbox, Tiny Honeypot, TCP Dump

6. CONCLUSION

The overall objective in this research is to investigate the needs of university computer labs in relation to security threats which plague them each day. Throughout the lifespan of this project, the goal is to allow Universities to become more information security aware and competent. By developing and implementing this technology (virtual honeypots), other research institutions will have a basis for continuous learning. As new virus threats are crafted each day, the only way to truly understand how to protect sensitive information and the machines that hold that information is to understand the nature of the threat and then select the appropriate action. The next step is to properly document the threat in such a way as to prepare the machines for that very same threat in the future.

7. ACKNOWLEDGMENTS

This work has been supported in part by U.S. Department of Education grant P120A080094 and by NSF CPATH.

8. REFERENCES

[1] Botnets – Security Focus, (Retrieved 5 September 2010). searchsecurity.techtarget.com

[2] Innella, Paul, (2001). The Evolution of Intrusion Detection Systems http://www.securityfocus.com/infocus/1514

[3] Owen, Greg, (Retrieved 25, December, 2010). Analysis and Comparison of Red Hat Linux 6.2 Honeypots – SANS Institute, Computer Forensics http://computer-forensics.sans.org/community/papers/analysis-comparison-red-hat-linux-62-honeypots-lids-enabled-kernels_23