

Approaches to Insider Threat Mitigation

By Sara Matzner

matzner@arlut.utexas.edu

Introduction

The actions of a trusted employee who is inappropriately accessing data can involve sabotage of crucial systems or unauthorized release of sensitive information. This "insider" is a threat to the fundamental integrity and confidentiality of information critical to any enterprise. To commercial entities, the acts of a malicious insider are estimated to cost millions of dollars in compromised intellectual property each year. To the government, this inappropriate access can be a threat to national security.

The external threat, the hacker who attacks a system from the Internet, is actively being addressed by technology. Network Intrusion Detection Systems (IDS) have evolved from basic research projects to the current array of sophisticated commercial products. However, transitioning solutions designed to combat barrier penetrations to solutions for the insider threat has a long way to go. In fact, it is feared that the insider threat will be even more difficult to mitigate than the external hacker.

While a few commercial products are now appearing that focus on the insider threat, many aspects of the problem remain largely unresolved. As with IDS, we need automated, labor-saving solutions. The solution is not to place IDS on the internal network, nor is it to simply use old technologies in a new way.

Currently, the burden of detecting insider activity rests with someone in the organization (a systems administrator or security officer) manually reviewing audit logs. That system administrator needs to be able to recognize suspicious activity, protect critical data, and monitor the behavior of the perpetrator to determine his motivation and ultimate goal. It would be ideal if he were able to do all these tasks using a toolset with a single interface that makes the complexity of the problem transparent to him.

Problem

The insider threat problem is different from that of the external hacker and the solutions will need to be different, too. The fundamental difference is that the insider is *supposed* to be on the system. He is an authorized, trusted user. He has privileges and perhaps intimate knowledge of the system. He may have physical as well as cyber access. He may be aware of protective measures that the organization has activated, and he may know the vulnerabilities of these protections. He may even be able to alter or delete the logs that contain evidence of his activities. The insider could be an employee, a contractor, or a service provider.

The tools developed for the insider threat have different requirements from those of IDS. The large numbers of false alarms that are produced by many of the current IDS products are completely unacceptable in insider threat detection. The resources that are called into play when there is evidence of inappropriate access to sensitive data are very expensive. Investigations of internal misuse can involve specialized taskforces of auditors, inspectors and, for government, counterintelligence officers. Therefore it is critical that the end result of any software tool designed to detect a

malicious insider is a very small list of names that, with a high level of confidence, identifies users requiring additional scrutiny.

Researching Solutions

The development of insider threat tools is taking a similar path to that of IDS. Some companies have released insider detection products early. At the same time, government agencies are taking an active role in sponsoring research programs to find more advanced solutions.

The Advanced Research and Development Activity (ARDA) sponsors a number of research efforts to address the insider threat problem. Applied Research Laboratories, The University of Texas at Austin, (ARL:UT) is participating in two of these ARDA-sponsored programs—a workshop on Cyber Indications and Warnings, and one of the Information Assurance for the Intelligence Community (IAIC) projects. The overall goals of our ARDA work are to develop innovative approaches to reliably detecting the presence of a malicious insider and providing some clues to his identity.

ARDA funds projects to provide specialized information assurance solutions for the intelligence community (IC). The IC is a unique security environment where the integrity and availability of information is extremely critical. While the focus of the ARDA-sponsored research is on protecting secure facilities from sophisticated attacks, the results from the research should provide solutions applicable to other, more general situations as well.

This research on the insider threat involves several different approaches that attack the problem from different angles. Many of the researchers involved have taken lessons learned from developing IDS to this new, more difficult arena. The approaches described below and illustrated in Figure 1 are just some of the projects supported by ARDA.

User-Centric Approach

The hypothesis for the most common approach to the insider threat problem is: To build a good defense, you must know your enemy. In the user-centric approach, researchers develop analytical models of typical users in order to simulate the adversary and to get a basic understanding of his behavior and motivations.

In the Cyber Indications and Warnings workshop, one group of researchers built models of user behavior based on actual, documented cases of insiders, including some very notorious espionage cases. Taxonomies were developed based on the insider's tactics, knowledge of the system, privileges, access, motive, and degree of risk aversion, and these were matched against a user's observed behavior ["Insider Threat Challenge Workshop Final Report," Maybury 2004].

The user-centric approach adds considerable value to solving the problem, but it is not a complete answer. A significant disadvantage is that the user-centric approach is based on signatures of known malicious attacks from past exploits. This is similar to signature-based intrusion detection systems for the external threat. From our work on developing IDS, we've found that historical information is a good starting point but not sufficient. So how do we allow for new knowledge?

Honeypots, Honeytokens

Honeypots may help address the shortcomings of the user-centric approach. The honeypot concept was developed originally for an external hacker. Participants from the Honeynet project joined the Cyber Indications and Warnings workshop to see if specialized honeypots deployed on an internal network could catch a malicious insider.

Think of a honeypot as a target that has been specifically designed to entice malicious users. The target can be a server, a database, or any common resource within the organization; but in this case, the resource has no production value. Instead, the resource's value comes from its illicit use ["Honeypots: Catching the Insider Threat," Spitzer 2003].

The novel concept of a honeytokens was developed as part of the workshop. Honeytokens are an instance of a honeypot: a specific data point is used as the bait. The idea was to instrument a resource with an attractive piece of information such as a password or credit card number and then track the information as it makes its way through the system. We would be able to capture the user's actions as he goes after the honeytokens, analyze his tactics and discover his targets. The information we gain by evaluating the insider's interactions with honeypots and honeytokens helps us to profile the insider, helping to build a better model for the user-centric approach.

During the Cyber Indications and Warnings workshop, discussions about using a version of the Honeynet concept brought out a major concern: While honeypots are meant to have no production value themselves, they are accessible by any user, the legitimate user along with the illegitimate. In developing the honeypot, we have made it look very much like a real system. Most likely, we have seeded the honeypots with phony but realistic data. What if a legitimate user wanders into a honeypot or is attracted by a honeytokens? Could this dummy target result in our inadvertent monitoring of an innocent user? If so, we would have a false alarm. Or worse, could the honeypot cause incorrect information to propagate throughout the system? If legitimate users enter the honeypot and use the fictitious information as though it were correct, we might significantly impact the work of the organization.

Critical Data Approach

In another research project sponsored by ARDA, ARL:UT is attacking the problem of the malicious insider from another angle. In what we've called our Advanced Detection Techniques (ADT) project, we are using a data-centric approach that is specifically designed for those data files, devices, or applications that are considered "the crown jewels" of an organization. This approach goes to the heart of the insider threat problem:

- ▲ The insider is a trusted employee with access and privileges.
- ▲ It is only when he abuses those privileges that he becomes a threat.

- ▲ We can't reliably determine beforehand who among our trusted employees may become an insider.
- ▲ Perhaps we should concentrate on the critical information rather than the people.

The assumption in this approach is that the most critical data warrants the most extensive protection. So, we are designing specialized sensors, or *sentinels*, to guard highly sensitive information wherever it resides throughout the enterprise. Typical resources that might justify this additional level of protection include any file or device containing classified or sensitive information or restricted applications.

The sentinels concentrate on all attempted access, whether or not the access is authorized. This focus on access may sound like a typical access or document control mechanism. However, in our approach, *any* attempt to access these critical resources activates a sentinel. In a typical access control approach, an insider would have appropriate access permissions. The sentinel responds by logging all access attempts and possibly initiating monitoring of the user's cyber activities throughout the network.

To develop these sentinels, we are using a combination of technologies. We have experimented with machine learning, including classification (support vector machines, hidden Markov models, decision trees) and clustering algorithms to identify misuse and anomalous behavior. We've instrumented finer-grained auditing and scrutinized system activity down to the kernel calls. We've generalized a system that we created for the external threat to handle the insider problem. The Generic Signature Model (GSM), a hierarchical signature-based system, allows the specification of arbitrary attack abstractions that are capable of capturing a variety of specific attacks ["Data Reduction through Event Correlation Using the Generic Signature Model," Lofaso et al. 2002]. And to test whether our approaches are working, we've developed a library of insider exploits.

Data Fusion Approach

An IDS monitors packets of incoming data to distinguish normal from abnormal network traffic. Systems for detecting the insider threat can also take this approach by monitoring activity on the internal network for suspicious events. As part of the Cyber Indications and Warnings workshop, the data fusion research group took a bottom-up approach to the problem, discovering evidence of malicious activity by correlating and analyzing a variety of low-level data. Our approach was based

on the hypotheses that while it is possible that evidence of malicious insider behavior might come from a single action, other examples of malicious behavior can be detected from an accumulation of low-level data from a number of diverse sources.

Whether the organization is a commercial entity or a government agency, the monitoring systems in place produce electronic audit logs that could provide indications of malicious insider activity. Currently, much of this data is being ignored or, if examined at all, someone must access and analyze each log separately. Our group of researchers set out to prove that audit log data combined with other data sources is valuable, and can provide significant insight into an insider's cyber activities.

We matched system, application, and network intrusion detection logs with physical access logs. Logs were compared with the user's role in the organization and the user's access permissions. The result of this data fusion was a "watch list" of names—individuals whose activities were suspicious and therefore might require further scrutiny.

Conclusions

The following list from "The Insider Threat to U. S. Government Information Systems" [NSTIS-SAM INFOSEC/1-99] provides a good set of guidelines of how best to address the insider threat:

- ▲ Define and enforce limits on overt access to sensitive data.
- ▲ Hold individuals accountable for their actions by providing non-refutable records.
- ▲ Review the actions of individuals from audit logs.
- ▲ Prevent covert access by making security measures resistant to attack.
- ▲ Detect covert access to sensitive information and networks.
- ▲ Quickly perform damage assessment.

To the above list I would add:


- ▲ **Know the enemy:** "If you know your enemy and know yourself, your victory may not stand in doubt." [Sun Tzu, *The Art of War*, c. 500 B.C.] User-centric research aided by what we learn from honeypots is helping with the first part of Sun Tzu's dictum.
- ▲ **Know what really matters to you and in what way it is vulnerable:** This point addresses the second half of Tzu's dictum. We need to determine what is our most critical data and how to protect it.

▲ **Know what you know already:** Information about activity on our internal networks is already overly abundant. So much of the audit logging we do is never actually reviewed. We need to collect only useful information and then automate the review process. An important discovery from our experiments was that standard audit logs currently are not designed with security in mind. In the course of our study, we found several ways in which these logs could be significantly improved for security purposes.

Have we solved the insider threat problem? Not yet; but we have now experimented with several approaches to the problem and have begun to combine the best aspects of these approaches into a working solution.

Acknowledgement

The Advanced Research and Development Activity (ARDA) supported part of this work. However, the content of this article does not necessarily reflect the position or the policy of the Government and no official endorsement should be inferred.

This work was performed, in part, in support of the Northeast Regional Research Center (NRRRC) which is sponsored by the Advanced Research and Development Activity (ARDA), a U.S. Government entity which sponsors and promotes research of import to the Intelligence Community which includes but is not limited to the CIA, DIA, NSA, NIMA, and NRO. 

Sara Matzner is the Program Manager for the Cyber Information Assurance branch of the Signal and Information Sciences Laboratory of the Applied Research Laboratories, The University of Texas at Austin (ARL:UT). She received a B.A. from the University of Maryland and an M.A. from The University of Texas at Austin. She directs a group of scientists performing basic research and developing prototype systems for a number of federal government agencies. Sara is a subject matter expert in the areas of information assurance and insider threat detection and as such publishes technical papers on these subjects and acts as a trusted advisor to government. She has worked for ARL:UT for 18 years with a one-year research assignment at the Los Alamos National Laboratory. Sara is a member of Capitol of Texas Chapter of ISSA and has served as its vice-president. She can be reached at matzner@arlu.utexas.edu.

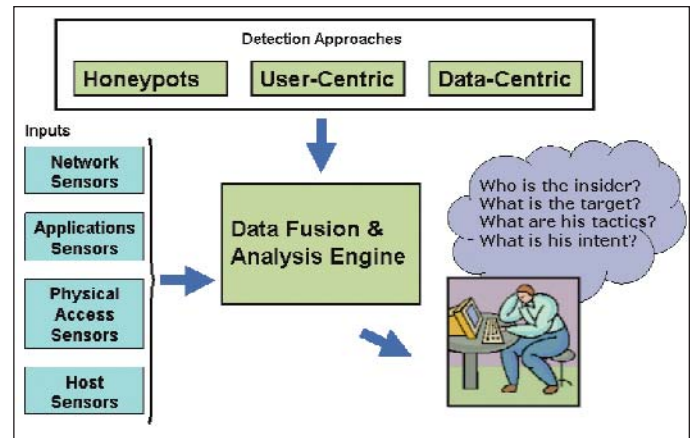


Figure 1: Convergence of Approaches