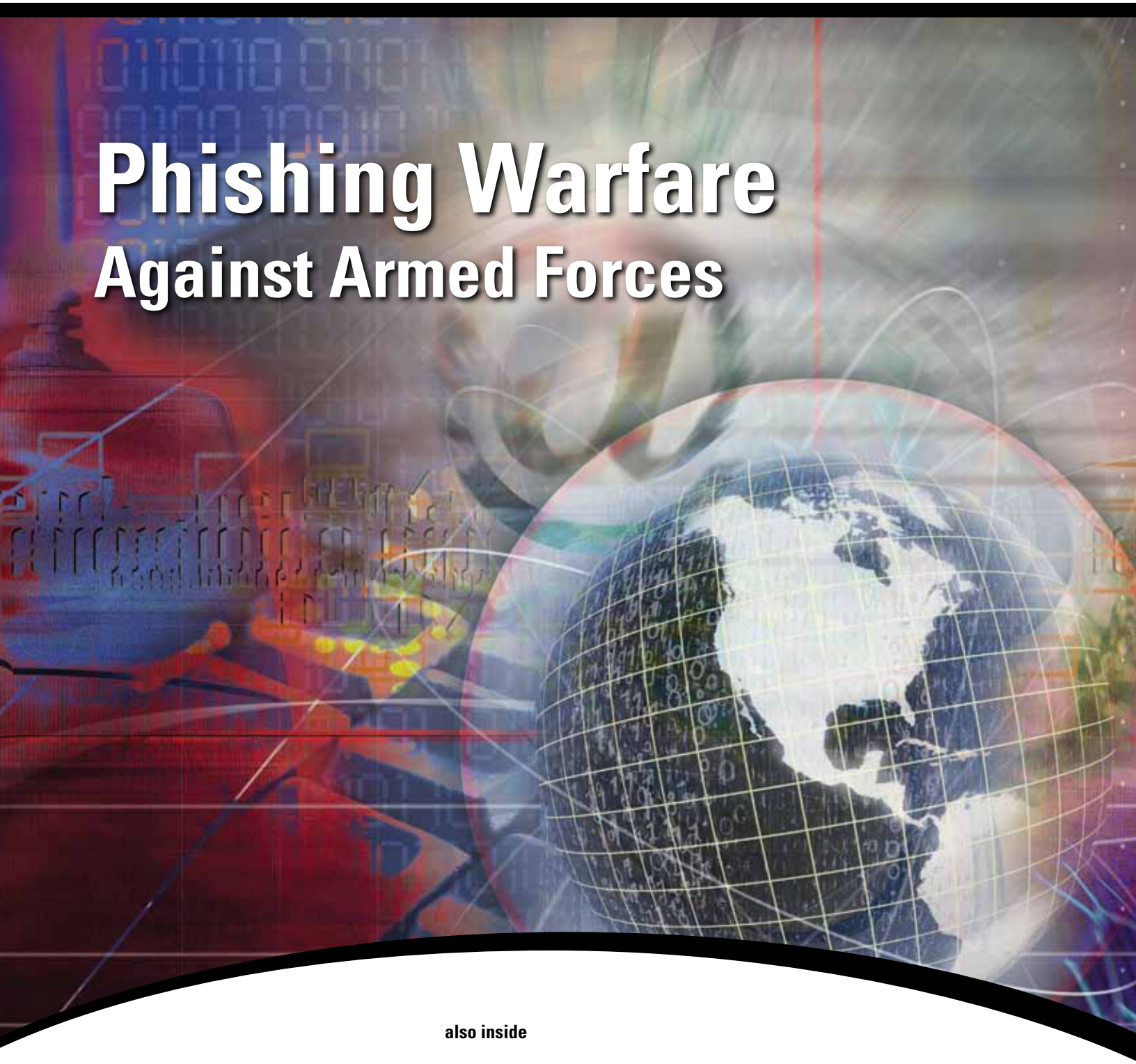
11/4

Volume 11 Number 4 • **Winter 2008**

IAnewsletter

The Newsletter for Information Assurance Technology Professionals



Phishing Warfare
Against Armed Forces

also inside

IATAC

contents

About IATAC and the *IAnewsletter*
The *IAnewsletter* is published quarterly by the Information Assurance Technology Analysis Center (IATAC). IATAC is a Department of Defense (DoD) sponsored Information Analysis Center, administratively managed by the Defense Technical Information Center (DTIC), and Director, Defense Research and Engineering (DDR&E).

Contents of the *IAnewsletter* are not necessarily the official views of or endorsed by the US Government, DoD, DTIC, or DDR&E. The mention of commercial products and/or does not imply endorsement by DoD or DDR&E.

Inquiries about IATAC capabilities, products, and services may be addressed to—

IATAC Director: Gene Tyler
Inquiry Services: Peggy O'Connor

***IAnewsletter* Staff**
Promotional & Creative Director: Christina P. McNemar
Art Director: Don Rowe
Copy Editors: Anne Roudabush
Designers: Michelle DePrenger, Dustin Hurt, Azi Pajouhesh
Editorial Board: Dr. Ronald Ritchey, Angela Orebaugh, Tara Bissett, Gene Tyler

***IAnewsletter* Article Submissions**
To submit your articles, notices, programs, or ideas for future issues, please visit http://iac.dtic.mil/iatac/IA_newsletter.html and download an "Article Instructions" packet.

***IAnewsletter* Address Changes/Additions/Deletions**
To change, add, or delete your mailing or email address (soft-copy receipt), please contact us at—

IATAC
Attn: Peggy O'Connor
13200 Woodland Park Road
Suite 6031
Herndon, VA 20171

Phone: 703/984-0775
Fax: 703/984-0773

email: iatac@dtic.mil
URL: http://iac.dtic.mil/iatac

Deadlines for future Issues
Summer 2009 January 14th, 2009

Cover design: Azi Pajouhesh
Newsletter design: Bryn Farrar, Donald Rowe

Distribution Statement A: Approved for public release; distribution is unlimited.

IATAC Chat

Gene Tyler, IATAC Director

On 02 September 2008, it was formally announced that Mr. John Palumbo would be stepping down as the ESSG Coordinator and Mr. Mike Frye would be assuming the position.

I can hardly believe we are already into the Fall season; time just seems to be flying past me again this year. One explanation for this could be from the many stimulating and new things going on in the IA community. From exciting conferences to emerging technologies, the IA arena is ever shifting. Another recent change in the IA community stemmed from the Enterprise-Wide Solutions Steering Group (ESSG). The ESSG was established to integrate and synchronize solutions, to advocate adherence to IA strategic goals, and to field enterprise-wide computer network defense (CND) solutions. The ESSG is co-chaired by United States Strategic Command (USSTRATCOM) and the Joint Task Force–Global Network Operations (JTF-GNO), with the ESSG Coordinator residing at USSTRATCOM. On 02 September 2008, it was formally announced that Mr. John Palumbo would be stepping down as the ESSG Coordinator and Mr. Mike Frye would be assuming the position. Mr. Frye has been serving USSTRATCOM now for close to 10 years, both as an active duty member of the US Air Force and also as a contractor. As the ESSG Coordinator, Mr. Frye will be ensuring the success of the ESSG's direction—which is to integrate and synchronize solutions for enterprise CND capabilities; advocate adherence to the Department of Defense's (DoD) IA strategic goals; establish enterprise-wide CND solution goals, objectives, and mission-based performance measures; and finally to support CND architecture development and migration strategy solutions. In order to accomplish this, he will have to help the ESSG voting members manage all aspects of the Group's activities, from "cradle to grave." In a discussion IATAC had with Mr. Frye, he stated, "There are a lot of major muscle movements that make up the ESSG activities, and the ESSG's success is solely dependent on 'detailed management' of those muscle movements. As the Coordinator, it is critical to keep those independent parts in motion and working in unison. I plan to continue to assist in coordinating this successful program that has been growing since its creation in 2002. I look forward to working with all the ESSG voting members and the rest of the ESSG team in the future." As coordinator, one of Mr. Frye's first responsibilities was to organize the September 2008 ESSG. This meeting was held at the Johns Hopkins Applied Physics Laboratory in Laurel, Maryland, from 30 September to 02 October. If you would like more information on the ESSG or the outcome of the September meeting, please do not hesitate to contact IATAC.

In this edition of the *IAnewsletter* you will once again, find some very interesting articles. We are once again privileged to have an article from AFIT entitled, "The EPOCHS Project: Creating a Framework for Managing Future Electric Power Systems in a Secure, Efficient, and Reliable Manner." We also feature a fascinating article on Phishing, "Phishing Warfare Against Armed Forces." This article reviews the idea that phishing of the armed forces could emerge as a new type of information warfare attack. In addition, we have an article written by two of our in-house Subject Matter Experts titled, "An Innovative Computer Forensic Technique for Recovering Deleted Files from Macintosh Computers." This article, as the title suggests, looks into how examiners are now able to recover deleted files from Macintosh computers. In addition to several other fascinating articles, we are also featuring our Spotlight articles. This edition focuses on the Department of Computer Science at James Madison University and Dr. Florian Buchholz.

If you have any questions, concerns, or recommendations for the *IAnewsletter*, please do not hesitate to contact me at *iatac@dtic.mil.* ■

Phishing Warfare Against Armed Forces

by Sean Price

Introduction

The problems with phishing persist. This scourge shows no sign of abating and will likely increase into the foreseeable future. [1] Warfighters are just as likely as any other group of individuals to be victims of a phishing attack. However, new types of phishing attacks may be used as surgical strike methods of exploiting weaknesses associated with phishing as an information warfare tactic. This article suggests that phishing Warfare Against Armed Forces (WAARF) will emerge as a new vector of information warfare.

Phishing has primarily been a phenomenon used to steal privacy information from unwitting victims. The principle attack method involves an enticement that causes the victim to visit a malicious website purported to be authentic. At its core, phishing is a high-tech form of social engineering. The attacker's objective is to convince the victim to disclose privacy information or credentials that the attacker can use to conduct a financial fraud. According to Myers, [2] a phishing attack is generally characterized by a lure, hook, and catch—

- **The Lure**—The lure is an enticement delivered through email. The email contains a message encouraging the recipient to follow an included hypertext link. The hyperlink often masks a spoofed uniform resource locator (URL) of a legitimate website.
- **The Hook**—The hook is a malicious website designed to look and feel like a legitimate website. The authentic-looking website asks the victim to disclose privacy-related information, such as user identification and password. Often the hook is an obfuscated URL that is very close to one the victim finds legitimate and is really a site under the attacker's control. [3]
- **The Catch**—The catch is when the originator of the phishing message uses the information collected from the hook to masquerade as the victim and conduct illegal financial transactions.

Phishing attacks are not new—the security community has known about them for some time. Unfortunately, people still succumb to these attacks. It is thought that a victim falls prey to phishing attacks because of the following reasons—

- **An email is considered authentic**; that is, a user is deceived into thinking an email is from an authentic or legitimate source. However, the email is often sent from an unrelated site, which can be gleaned from detailed header information in the email message. Unfortunately, email client software often hides this detailed information from the end user. Failure to verify this information can cause the user to believe the email is genuine when in fact it is not. The "From" field of the email is typically spoofed to appear to originate from a legitimate sender.
- **The request seems legitimate.** The recipient perceives the enticement of a phishing email to be valid.
- **The website appears genuine.** Clicking on a link in an email commonly spawns a new browser instance or causes the most recently activated instance to navigate to the associated URL.

Spear phishing is a directed type of attack that targets specific groups of people. With this attack, the phisher sends an email to group of people who are often in the same organization. Frequently, the phishing email is spoofed to appear to be from an actual member of the group. Phishing WAARF is considered a type of spear phishing. The important difference between the two is principally the objective of the attack. Spear phishing is conducted to perpetrate a fraud, whereas phishing WAARF is used to gain military intelligence, conduct espionage, or perform information warfare activities.

Phishing represents a potential attack vector for terrorists, nation states, and militaries. As opposed to financial gain, phishing could be used to steal credentials or compromise a host in a target network. An enemy could establish a back door into an unclassified network and obtain sensitive information that it

could use to compromise operational security. An enemy could use a back door in a sensitive system not only to gather intelligence but also to conduct information warfare activities, such as altering information or disrupting important systems during strategic events. Indeed, phishing is an ideal social engineering weapon that the attacker can leverage to take advantage of user trust, complacency, or ignorance.

The perception is that two possible attack vectors exist that armed forces could use against each other. One vector is considered more direct and the other relies on layers of deception. The following describes these potential attack vectors—

- **Obfuscated Phishing WAARF**—This attack is similar to an attack ordinarily used by a criminal seeking financial gain. The source of the attack is obvious by looking at the indicators, such as email headers and website addresses. At its best, an obfuscated phishing WAARF is a camouflaged supported attack.
- **Covert Phishing WAARF**—In this attack, the attacker uses various levels of deception, misdirection, and system abuse to achieve the objective with as little evidence pointing back to the attacker as possible. A well-planned and executed covert phishing WAARF represents a truly clandestine operation.

Conjectural Multi-Vector Phishing WAARF Attacks

Suppose we have two factions, the Blue Union and the Red Menace, who are engaged in information warfare. The Blue Union relies heavily on its cyber infrastructures for command and control operations. Members of the Blue Union often communicate through email and share information through websites operated by each base of operation. The Red Menace seeks to disrupt Blue Union operations using phishing WAARF attacks.

Figure 1 depicts a classic example of a phishing attack. In this scenario, which epitomizes an obfuscated phishing WAARF, the Red Menace sends an email to an unsuspecting Blue Union warfighter. The link in the phishing email sends the warfighter to a Red Menace-controlled server that is nearly identical to the spoofed public server. In this situation, the Red Menace may collect privacy-related financial information or any other information that gives insight into the activities or character of the warfighter. The objective of the Red Menace is not to conduct a fraud but rather to gather intelligence about the warfighter and learn operational aspects of the Blue Union forces. If the Red Menace is able to collect enough information about a large number of Blue Union warfighters, it may be possible to predict operational activities, such as deployments, command changes, and other operationally sensitive information.

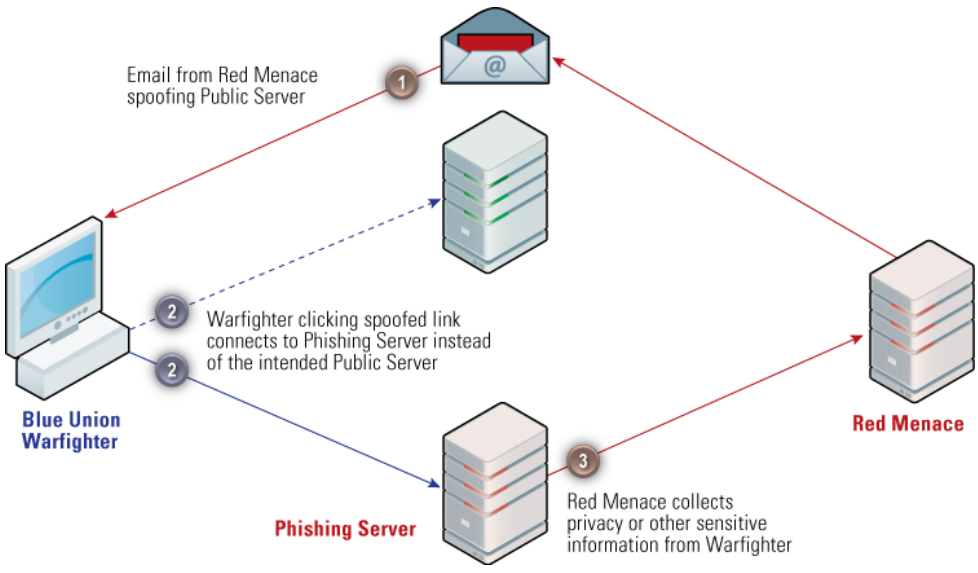


Figure 1 Classic Phishing Attack

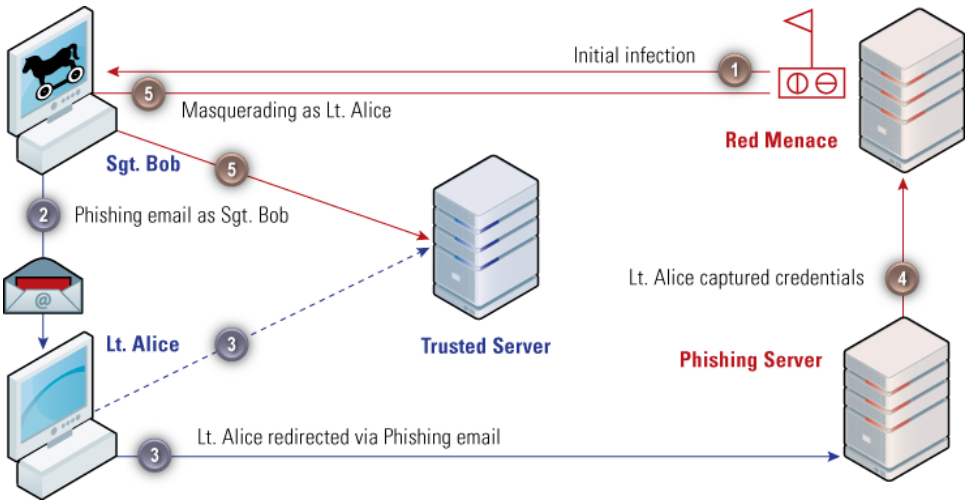


Figure 2 Trojan Phishing

This first scenario has an even darker side. The collection of privacy-related information can also be used against individual warfighters. For instance, if a particular warfighter has financial problems or is participating in dishonorable activities, the Red Menace might use this information to entice or blackmail the warfighter to disclose sensitive or even classified information. Indeed, the collection of personal information could be used as an impetus for espionage or intelligence-gathering activities by targeting susceptible warfighters.

Phishing can also be used as a secondary vector to compromise a server. In Figure 2, a compromised workstation is used in an effort to steal the credentials of another user. The Red Menace has managed to load a Trojan horse onto Sgt. Bob's workstation. The Trojan enables the Red Menace to operate in the security context of Sgt. Bob. The attack involves the transmission of a legitimate email from Sgt. Bob to Lt. Alice. The Red Menace sends an email through the Blue Union email system using the email client or through appropriate protocols, such as Simple Mail Transfer Protocol (SMTP) or Messaging Application Programming Interface (MAPI). Lt. Alice receives the email from Sgt. Bob with a spoofed link to a trusted server. Lt. Alice trusts the email because Sgt. Bob has sent it, but she is unaware that the link is spoofed and actually connects to an untrusted server controlled by the Red Menace. If the spoofed server has a login interface, the Red Menace can capture the credentials and subsequently use them to gain access to the trusted server *via* the Trojan horse on Sgt. Bob's workstation.

In this scenario, a Trojan compromise is combined with the social engineering tactics of a phishing attack to gain further access into a system. This compromise permits the Red Menace to monitor or modify the activities of Lt. Alice on the trusted server. The devious nature of this type of attack easily classifies itself as a covert phishing WAARF activity.

Attackers can also use flaws in web servers to steal information from an unsuspecting warfighter. Figure 3 shows two possible cases of this scenario that involve the exploitation of a cross-site scripting flaw in a trusted server. This type of vulnerability enables an attacker to redirect browser communications from the trusted server to another server under the control of the enemy. The attacker can accomplish the first method of exploitation by directly inserting code into the affected server. In this case, a browser requests web content with a malicious script that makes an unintended connection with another server. The second case occurs when a user clicks on a link in an email containing an embedded script that compromises the trusted server.

In both situations, when the user connects to the trusted server, the cross-site scripting attack opens a connection from the warfighter's browser to another server. The browser of the affected Blue Union warfighter might send information to or download content from the server under the control of the Red Menace. This act provides the Red Menace with the ability to manipulate browsing activity on the affected trusted server or collect intelligence information about Blue Union activities. The use of cross-site scripting flaws is not new [4] and more overt. As such, this type of attack is considered an obfuscated phishing WAARF.

An examination of these scenarios reveals a commonality of weaknesses. These scenarios exemplify the following persistent issues—

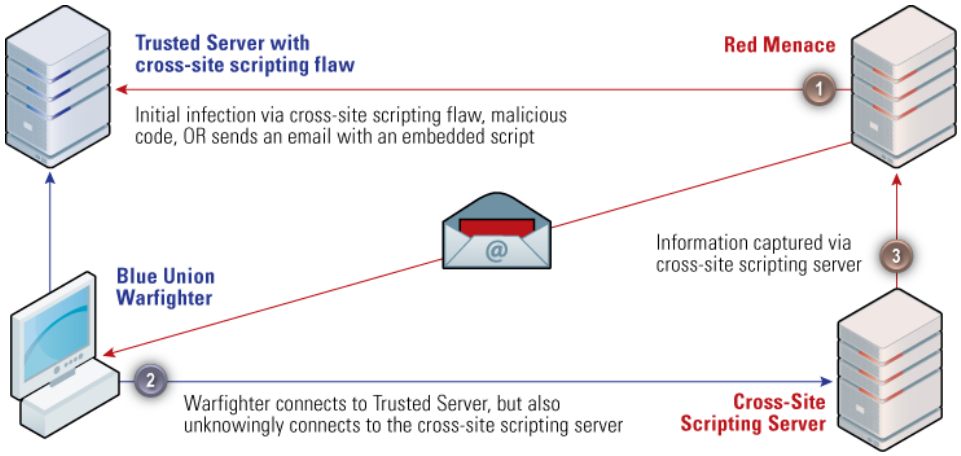


Figure 3 Phishing With Cross-Site Scripting Attacks

- Users do not sufficiently check the integrity of email messages
- Users are often unaware of actual browser connections
- Enforcement measures are lacking.

Prior Efforts

Some of the earliest security tools developed to fight phishing attacks include specialized security toolbars running in the browser. These toolbars either rely on black lists of known phishing sites or identify attributes of a loaded webpage that seem suspicious. Some researchers have concluded that these types of tools are not always effective because end users ignore passive warning indicators. [5] However, users do seem to pay attention to more interactive indicators.

One line of thought is to interrupt phishing emails before users receive them. SiteWatcher is a system that scans incoming email for phishing indicators [6]. This experimental tool operates by extracting email-embedded URLs, downloading content from the location, and comparing it with internal copies of protected webpages. If a high degree of visual similarity is found and the URL does not match the protected webpage, SiteWatcher issues an alarm.

Another possible approach is to evaluate hyperlinks in emails when the user clicks them. An experimental tool recently created compares the latest browser instance URL with a list of protected Internet host names. [7] If a URL is found to contain aspects of the protected URL, then the browser instance is closed and a new one is opened on a protected desktop. This provides the user with a strong visual indication that a clicked link contains a protected site and further action has been initiated to protect the browsing activity.

A New Technology Defensive Method

This article proposes a technical countermeasure that could be used to protect the warfighter from accidentally succumbing to a cleverly devised phishing WAARF. We note that the previously described phishing WAARF scenarios identify direct and indirect attack methods. Attributes these attacks have in common are—

- An email from an apparent or actual trusted source
- A connection with a spoofed website outside of the trusted IP range or Internet domain.

Our countermeasure accepts the possibility that the first attribute will not be avoided. Although a spoofed email contains telltale signs of its illegitimacy, it is more difficult for users to determine if an email is from an actual end user or was generated through a Trojan horse executing in their context. The proposed solution instead will provide users with browsing activity status and will enforce a security policy by—

- Monitoring browser connections
- Providing visual indications to the end user regarding browser site access
- Enforcing policy that excludes concurrent browser instances with trusted and untrusted sites.

The goals of our solution are to identify activities that are phishy and enforce policy compliance. There are two technical ways to accomplish this when Internet Explorer is the browser used. One method involves the use of an in-processing monitoring component, such as Browser Helper Object (BHO). This method has the following advantages—

- Monitors activity for each instance
- Has full access to all objects in the browser
- Can take actions based on browser activity.

The second method involves a separate process that monitors the activity of each browser instance through Windows Shell and automation techniques. This method could evaluate the URL in each browser as well connections with sites outside the base URL in the browser's address. This method has two other desirable attributes. First, browser flaws do not immediately affect it, which helps ensure the security mechanism will continue to function even when a browser is compromised. Second, in environments that do not allow the use of BHOs for security reasons, the use of an external monitoring process is a more agreeable approach. An additional consideration involves coordination between BHOs for multiple instances of Internet Explorer. This coordination may necessitate complex intra-process communications that could be problematic. In this regard, a browser vulnerability is less likely to affect the second method suggested while achieving the same goals as the first method.

Tool Design

An experimental tool, called the Browsing Activity Authenticity Tool (BAAT), was developed to identify phishing activity and enforce a defined policy. BAAT runs in the context of the user and monitors browser activity from two vantage points. The first perspective involves enumerating Internet Explorer object collections made available by the operating system. This provides the capability to identify the URL information for each browser instance. It also enables the BAAT to terminate browser instances that violate policy or demonstrate suspicious behavior. The second observation point involves monitoring all network-related activity. This is similar to running the netstat.exe tool to evaluate network activity of executing processes. Each browser instance is monitored for any connections that are contrary to policy, allowing for the identification of an authorized site that may redirect the browser to download content from a questionable location. Again, this enables the BAAT to monitor for policy violations or enforce compliance.

There are three primary aspects to the tool's operation: a browsing policy, indicators, and enforcement rules. A browsing policy identifies authorized locations, unauthorized locations, and any others of an arbitrary nature. Indicators are available as an icon in the Windows System Tray and provide the

user with a status of the actions taken by the browser with respect to the identified policy. Finally, enforcement rules provide a capability to take automated actions based on policy deviations experienced by any particular browser instance.

Browsing Policy

The policy comprises five items that specify the handling of site addresses and URL composition—

- **White List**—A listing of protected addresses and sites that have no restrictions; these locations are explicitly trusted for connections with unknown locations.
- **Gray List**—Sites that are trusted but are not allowed to redirect the browser to download content from unknown sites.
- **Black List**—Prohibited addresses and sites; at no time should a browser download content from these sites.
- **Unknown Sites**—Any location not explicitly identified in any of the lists.
- **Spoofing—**A host aspect of the URL contains other domain information as part of its address (*e.g., http://www.faq.com.bad.net* is considered a spoofed address because bad.net appears to be spoofing faq.com); finding domain extensions in the host name is considered highly irregular and indicative of spoofing activity commonly experienced with phishing attacks.

Visual Indicators

Each indicator provides a status of the collective policy compliance of all browser instances. The indicators are visible as icons in the Windows System Tray. Icon meanings are as follows—

- **Steady Green**—All browser instances are communicating with white- and gray-listed sites.
- **Flashing Green and Yellow With a Balloon Popup**—A gray site download content from an unknown location; the balloon popup provides an additional notification of the offending location for extremely short-term connections.
- **Steady Yellow**—An unknown site is open in a browser instance.
- **Flashing Yellow and Red**—An unknown site appears to be spoofing another.
- **Steady Red**—A connection has been made to a black-listed site.
- **Flashing Red—**A browser location is spoofing a gray–or white–listed site.

Enforcement Rules

The BAAT monitoring tool can follow up on enforcement rule violations through various configured actions, including—

- Terminating instances that connect to black-listed locations
- Notifying the user and terminating browser instances apparently spoofing a white-listed site or range
- Notifying the user and terminating browser instances spoofing any site
- Providing indications only.

> Trust will become more difficult as websites mesh with others through Web 2.0. Those charged with defending systems must bear in mind the complexity and dynamic nature emerging in online systems.

Experimental Setup

The experiment was designed to test the ability of the tool to detect spoofing activity and enforce the policy. Several test URLs were either opened through an email link or entered into the browser's address bar. The test URLs were categorized as acceptable or suspicious. Subsets of the suspicious URLs are those that—

- Spoof a protected website host name by including it in the host name of another
- Include spoofing activity in the path or search part of a URL
- Contain downloads from a black-listed website either directly through the browser address bar or by connections made through a target webpage.

Results and Future Work

The application performed acceptably in all areas of the experiment. The results of this preliminary effort are encouraging, but more work is needed. Human computer interface evaluations should be conducted. It is important to know if users will find this type of tool beneficial or acceptable. Will users pay attention to the phishing indicators or will they ignore the indicators? Users sometimes are resistant to strong enforcement mechanisms. A study of user attitudes and actions regarding the various levels of enforcement rules will help determine the viability of this solution from the user's perspective. A deeper investigation into the robustness of this approach is needed. It is unknown at the moment whether complex attacks can be used to circumvent this type of monitoring. This factor is important to consider given the constant escalation of information security attacks and attempted countermeasures.

Discussion

This approach seems to be a straightforward implementation. Unfortunately, this is not the case. Websites are becoming very complex networked applications. Many sites share information and rely on passing information from one to the other *via* browser interactions. Information sharing

is becoming more the case as the adoption of Web 2.0 technologies increases. Trust will become more difficult as websites mesh with others through Web 2.0. Those charged with defending systems must bear in mind the complexity and dynamic nature emerging in online systems. The experimental prototype is an endpoint tool, the need for which is becoming more evident over time. [1] Careful consideration and evaluation is needed when instituting controls to protect systems. Similarly, risk should not be accepted just because the task of technology implementation is formidable.

Phishing WAARF Defense

The previously described tool is just one technical approach to help combat the threat of phishing WAARF. Real security is not a product, but rather a process. Indeed, information assurance encompasses multiple aspects and layering of processes to defend systems and their information from an assault. It is advisable to follow frameworks and models that mesh security controls as the basis for a security program. The Information Assurance Model is one approach that identifies people, operations, and technology as the categories of security countermeasures. [8] Implementing controls in each of these areas supports a defense-in-depth strategy recommended by the National Security Agency. [9]

People

System users are the core of any security solution. Individuals must understand associated threats, risks, and countermeasures and must receive sufficient guidance and training to react appropriately when exposed to a phishing attack.

- **Training**—Educate personnel about phishing WAARF threats. Clearly communicate processes and procedures to be used to defend against such threats. Train users on the proper use of tools deployed to defend against these types of threats.
- **Evaluation—**Evaluate a sample of system users periodically to determine the effectiveness of the training. Subject users to an organizational-generated phishing attack that is as lifelike as possible. Protect any information revealed by the users from unintended exposure. Provide remedial training to users who inadvertently succumb to the phishing scenario.

Operations

Security policies, strategies, and tactics help defend a system against attack. Operational countermeasures rely heavily on explicitly documented and repeatable processes that lend themselves to evaluation.

- **Policy**—Promulgate policy that clearly identifies what is considered phishing and the controls that should be in place to defend against these types of attacks. Set policies that establish the strategic basis for combating phishing attacks.
- **Design**—Select countermeasures that support policy. Identify aspects of the system that might increase the likelihood of a successful phishing attack and develop compensating controls where appropriate. Consider the implications of a successful phishing attack. Integrate detection methods and response actions into monitoring mechanisms that would be enacted when other controls fail. Preplanned responses are essentially an additional defense in depth layer, so ensure the design is properly documented. A security design supports the battle plan used to defend against phishing attacks.
- **Procedures**—Fully document all procedures associated with the design and implementation aspects of defending against phishing attacks. These necessary documents provide the tactical response needed to defend a system against attack. Clearly identify system and application configurations in appropriate documentation. Ensure user documentation regarding the use of tools and procedures to undertake when a compromise is suspected is clear, concise, and readily available. Document procedures for security personnel, administrations, and operational support staff regarding phishing attacks, and make them accessible to those with operational duties.
- **Assessment—**Because an assessment provides an information assurance backbone for the system and the organization, periodically evaluate the effectiveness of phishing countermeasures. Ensure the assessment determines whether the controls are properly functioning, operating as intended, and producing the desired output. Redress policy, design, and procedures when weaknesses are identified.

Technology

Technology enables automated mechanisms to enforce a system's security policy. A number of technical measures can be used to prevent and detect phishing attacks—

- **Disable HTML in Email**—Disable HTML to remove any obfuscation attempts to hide the actual URL of a hyperlink using an HTML link title.
- **Validate or Authenticate Email Messages**—Use email servers as well as client-side tools to detect bogus email header information. Spam filters are quite useful in this regard. Either discard the message or inform end users that the email contains suspicious attributes.
- **Restrict Browser Capabilities**—Increase the security levels of the browser to prevent the download or execution of untrusted mobile code that the browser cannot constrain. For instance, ActiveX and Flash content run as all or nothing and can interact with any file or resource accessible to the user. Allowing the execution of unapproved

content is not only risky but also essentially violates change control management for the workstation.

- **Detect Network Intrusion**—Detect connections with black-listed sites through network intrusion detection. Correlation between intrusion detection and email logs can be conducted to identify additional websites that should be black listed or other users who may have also received similar messages.
- **Detect Host-Based Intrusion**—Use endpoint intrusion detection to monitor for covert phishing WAARF, such as a Trojan Horse interacting with an email client or sending email *via* SMTP. These tools can also be used to detect phishing sites that might attempt to exploit a flaw in the browser.
- **Secure the Organization's Domain Name Service (DNS)**—Ensure appropriate controls are established to prevent the tampering of local DNS servers. Clients depend on the accuracy of the information in the DNS. Breaches in the integrity .of the DNS server can have serious consequences on an entire organization.
- **Secure Workstation Host Files**—If it is not necessary to modify this file, set the access control lists to Read for all users. This can help prevent modification of the file, which could cause the browser to point to a malicious site.
- **Use Defensive Anti-Phishing to Combat Fake Sites**—Investigate and implement tools specifically designed to respond to phishing activity detected on workstations and servers. These types of tools can be used to enforce policy rules that can disrupt overt and covert phishing WAARF attacks.
- **Provide the User With Visual Indicators of Questionable Browser Activity**—Implement additional tools that monitor browser activity for security violations, anomalous behavior, or connections with known or suspected phishing sites. These tools should, at a minimum, provide a visual message or event to alert the user to activity that might be suspicious or dangerous.
- **Close Offending Browser Instances When Violations Occur—**Use an automated tool, such as the BAAT, to immediately terminate browser instances violating a policy. This helps protect the integrity of the workstation and user information. Furthermore, this action is another visual indicator for the end user that a suspicious connection has been made through the browser.

Conclusion

Phishing attacks are not simply a problem facing the public—phishing attacks will likely focus on military targets in the future. System owners, administrators, and security professionals must design automated controls into systems to detect and prevent overt and covert phishing WAARF attacks. Enforcement mechanisms that automatically react when a policy is violated will likely yield the best security results. However, no one tool should be relied on exclusively. The proper application of defense-in-depth measures involving people, operations, and technology will be needed to counter these types of attacks as they emerge in the future, if they have not already occurred. ■

References

1. Kuper, P. (2006). A Warning to Industry—Fix it or Lose it. *IEEE Security & Privacy,* 4(2), 56–60.
2. Myers, S. (2007). Introduction to Phishing. *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft.* (eds. Jakobsson, Myers). Hoboken, NY: John Wiley & Sons.
3. Fitzgerald, T. (2008). The Ocean is Full of Phish. *Information Security Management Handbook,* 6th ed. Vol. 2. (eds. Tipton, Krause). Boca Raton, FL: Auerbach.
4. McRee, R. (2008). Anatomy of an XSS Attack. *The ISSA Journal,* 6(6), 12–14.
5. Wu, M., Miller, R.C., and Garfinkel, S.L. (2006). Do security toolbars actually prevent phishing attacks? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 601–610.
6. Liu, W., Deng, X., Huang, G., and Fu, A.Y. (2006). An Antiphishing Strategy Based on Visual Similarity Assessment. *IEEE Internet Computing,* 10(2), 58–65.
7. Price, S. (2007). Protecting Privacy Credentials from Spyware and Phishing Attacks. *Proceedings of the IEEE SMC Information Assurance and Security Workshop,* 167–174.
8. Maconachy, W.V., Schou, C.D., Ragsdale, D., and Welch, D. (2001). A Model for Information Assurance: An Integrated Approach. *Proceedings of the IEEE SMC Information Assurance and Security Workshop,* 306–310.
9. National Security Agency (2008). Defense in Depth. Retrieved June 15, 2008 from *http://www.nsa.gov/snac/support/defenseindepth.pdf.*

About the Author

Sean M. Price, CISA, CISSP | is an independent security researcher and consultant living in northern Virginia. He specializes in designing and evaluating organizational information assurance programs and system security architectures. His research interests include insider threat, information flows, and applications of artificial intelligence to information assurance problems. His prior publications include the *Information Security Management Handbook, Official (ISC)²Guide to the CISSP CBK,* and *IEEE Computer* magazine, as well as other journals and conferences. You can email him at *sean.price@sentinel-consulting.com.*

Dr. Florian Buchholz

by Angela Orebaugh

This article continues our profile series of members of the Information Assurance Technology Analysis Center (IATAC) Subject Matter Expert (SME) program. The SME profiled in this article Dr. Florian Buchholz, Assistant Professor in the Department of Computer Science at James Madison University. Dr. Buchholz's research interests are in cyber-forensics, network security, and operating systems security. He is currently involved in two significant research efforts—timestamp correlation and data carving.

Timestamp Correlation

When analyzing digital evidence during a computer forensics investigation, an investigator frequently utilizes time information to determine the proper sequence of events or to relate events on a computing system to those in the real world. For this purpose the investigator uses timestamp data (timed events) from various sources of digital evidence such as log files, file system metadata, and applications. However, there are several problems with the way timestamps are generated and stored on computing systems, which leads us to question the accuracy and reliability of the time information they present. Computer clocks that generate the timestamps may not be set to the correct time for the following reasons—

- Clock skew
- Setting of the clock to an arbitrary time
- Precision of the clock.

Timestamp information is subject to—

- Rounding
- Loss of precision when storing the information
- Tampering

These considerations play a role when trying to tie computing events to particular real world times, and the problem is considerably magnified when dealing with evidence from disparate computing systems, each with different clocks and timestamp behavior. In the latter case, an investigator faces the added challenge of establishing the correct order of all events among the different systems.

Dr. Buchholz is performing research to provide a framework for forensic investigators to correlate timestamps to a common reference time, *e.g.* Universal Coordinated Time (UTC). To accomplish this the investigator will need to translate all timestamps from a given system into the domain of the reference time. Furthermore, to address rounding behavior and different timestamp precisions, the time for an event in the framework is not represented by a discrete time, but rather a time interval. Comparing times for events then becomes a matter of evaluating the overlap between given time intervals.

To translate a timestamp to a reference time the investigator needs a complete history of the computer clock that created the timestamp. While past research has shown that clock skew for a computing system is mostly constant (*i.e.,* the clock drifts away from real time in a linear fashion), the complete history is required to be aware of all the times a particular clock on a computer was modified (either manually by a user or through service programs on the system). In general, however, this information is not readily available to the forensic investigator. Therefore the framework also should assist the investigator in identifying any past changes in the clock's value. Dr. Buchholz is currently exploring techniques that do just this. Obvious candidates are jumps "back in time" in ordered log data, "missing" regular system activity from the logs (they could indicate a clock jump forward in time), and the building and evaluating of a happened-before-relation graph of events on the system. Any time information associated with the events in the graph should be consistent with the happened-before-relation, or some sort of clock (or timestamp) tampering is likely to have occurred. Also, cycles in such a graph should not occur unless there are inconsistencies with the evidence due to tampering, clock manipulations, or incomplete evidence.

An investigator faces an added problem whenever a clock is set back in time. Just by looking at a timestamp with a value that was recorded during the time

period that the clock runs through again, it is impossible to determine whether the timestamp was recorded before or after the jump back in time. For example, if a clock that was representing UTC was set back 1 hour at 5:00 pm (clock now shows 4:00 pm and is one hour behind UTC), a timestamp with a time of 4:30 pm system time could be associated with 4:30 pm UTC as well as 5:30 UTC. The framework will need to provide some aid to the investigator to resolve this ambiguity. For example, if there are other events with known exact reference times that the investigator can link to the event, he or she can possibly make a decision for the original event.

The overall goal of the framework is to create an accurate unified timeline for events from disparate sources in any given reference time. Further research should then investigate visualization and presentation techniques to increase the usefulness of such data in forensic analyses and presentation to non-technical people.

The overall goal of the framework is to create an accurate unified timeline for events from disparate sources in any given reference time.

Data Carving

The discipline of data carving is concerned with the retrieval of files (or partial files) from a storage medium where little or no file system information is present, as well as the retrieval of information from parts of the storage not currently allocated by a file system (*e.g.*, slack space, unused blocks, or unused partitions). Under the assumption that no file system information is present, data carving usually takes 512 byte disk sectors as the smallest building block in the attempt to reassemble files.

A very simple carving method is called "header/footer carving," where the assumption is that the file is stored on the disk in consecutive sectors. Once a well-known file type header is encountered at the start of a sector, all subsequent sectors are "carved out" until a corresponding footer is found (or some threshold is reached). The resulting file is then verified whether or not it is a correct file. This technique may yield a large number of files from a given storage medium, but due to fragmentation, missing information, or file types that do not have easily identifiable headers and footers other files may be missed. Thus current research focuses on those scenarios where header/footer carving will not yield results.

An early approach to address the shortcoming of header/footer carving was to use a combinatorial approach where all combinations of remaining sectors were evaluated for correct files. However, depending on how many sectors need to be analyzed, the complexity of this approach does not scale very well. Promising recent approaches are using statistical information on each sector's data to determine the sector's file type or to which particular file a given sector may belong. These approaches work very well for certain file types and when sectors belonging to the same file are still within close proximity of each other.

Dr. Buchholz's file reassembly research involves a classification system of disk sectors that are applied before the actual carving step. The classification data can then speed up the carving algorithms already in existence, or may lead to new techniques altogether. The classification system looks at the sector data based on a particular file type and determines if the sector starts with, ends with, or contains data structures consistent with the file type. This should also take partial information into account for the start and end cases. The goal is to "describe" the sector like a puzzle piece and then query which other sectors can be adjacent puzzle pieces. For example, let's assume we have one 512 byte disk sector that starts with 200 bytes of unknown data (seemingly random), followed by a PKZip local file header (let's assume this is 40 bytes long) followed by 272 bytes of unknown data. The PKZip header may reveal that the length of this entry is 1000 bytes. In this example we can conclude the following—

1. The first sector following this sector must not contain any PKZip headers, but rather 512 bytes of compressed data (high entropy).
2. The second sector following this sector must contain 216 bytes of compressed data followed by another PKZip header (either another local file header, a data descriptor, an archive header, or a central directory header).
3. The sector preceding this sector must either be a sector containing only compressed data or contain a PKZip header in such a way that it fits together with the 200 initial bytes of this sector.

If such a query can be done efficiently, this will reduce the search space that the existing combinatorial and statistical carving methods must use, and could significantly improve their performance. This approach will most likely work best with well structured data formats (such as zip files, JPG files, HTML files, *etc.*), but can probably also be extended for data structures used by certain compression algorithms (PKZip's deflate or JPG) to obtain a more fine-grained classification. ■

References

More information on Dr. Buchholz and his research and publications may be found at *https://users.cs.jmu.edu/buchhofp/*.

Information Security at James Madison University

by Angela Orebaugh

James Madison University (JMU), founded in 1908, is centered in the beautiful Shenandoah Valley, Virginia. JMU offers more than 100 degree programs on the bachelor's, master's, educational specialist, and doctoral levels and is home to over 17,000 students. The College of Integrated Science and Technology (CISAT) houses the Computer Science department, which offers a graduate degrees in Information Security (InfoSec) and Secure Software Systems (SSS).

The JMU InfoSec program was established in January 1997 as one of the first graduate Information Security programs in the nation. It is also one of the original seven NSA designated Centers of Academic Excellence in Information Assurance Education (CAE/IAE). In addition to a Master of Science in Computer Science degree, all Infosec graduates receive two NSA approved certificates: Information Systems Security (InfoSec) Professionals (NSTISSI No. 4011) and Information Systems Security Officers (CNSSI No. 4014). The InfoSec program is designed for working professionals and 1999 the program moved to 100% distance learning. Students connect to the Internet and access courses, discussion forums, realistic online labs, and virtual classrooms hosted on the Blackboard Learning System. The program is cohort-based and requires 33 credit hours of graduate coursework. Students may choose to complete a thesis or take a comprehensive exam. The InfoSec program is led by Dr. Hossain Heydari and includes faculty with diverse backgrounds and research interests such as computer, network, and operating system security, fault tolerant systems, digital forensics, multicast security, intrusion detection, vulnerability assessment, secure software engineering, cryptography, and privacy.

The JMU InfoSec program was established in January 1997 as one of the first graduate Information Security programs in the nation.

JMU also offers an Information Security Masters of Business Administration (MBA) degree, with a goal of creating executive managers who understand the business implications of information security. The InfoSec MBA is a joint effort between the CISAT InfoSec program and the College of Business. The InfoSec MBA focuses on managerial decision-making, analytical problem solving, oral and written communication, and application of theoretical constructs all set in an InfoSec framework. Graduates of this program will understand preservation of information confidentiality and protection, risk management, data and system integrity, availability, authenticity and utility all set against a strategic business background. The InfoSec MBA degree is a distance-learning based program using both Blackboard and audio and video streaming technologies.

The Secure Software Systems program focuses on software engineering practices that include consideration of security in design, construction, testing, and deployment. Students learn practical techniques and tools for secure software development. The SSS program includes 36 credit hours in Computer Science, Software Engineering, Information Security, and elective or thesis hours. Students have the option to complete a thesis, however all students must pass a comprehensive exam after the first year of study. ■

References

1. For more information on JMU's InfoSec degree please refer to *http://www.infosec.jmu.edu*
2. For more information in JMU's InfoSec MBA please refer to *http://www.jmu.edu/mba/aboutinfosec.htm*
3. For more information on JMU's Secure Software Systems degree please refer to *http://www.cs.jmu.edu/sss/index.html*

An Innovative Computer Forensic Technique for Recovering Deleted Files from Macintosh Computers

by Aaron Burghardt and Adam Feldman

Recovering deleted files is a critical technique employed frequently in computer forensic examinations. Such files often contain residual information that may be relevant to investigating computer-aided criminal activity, network intrusions, and computer network attacks. The recovery potential of deleted data and files varies based on the host operating system filing system mechanisms that pertain to deleted files and deleted blocks. Macintosh computers are increasingly seized/acquired for computer forensic examinations, however current deleted file recovery tools and techniques for Macintosh computers lack sophistication and accuracy.

Current File Recovery Approaches

In general, deleted file recovery follows one of two basic approaches. The first approach involves scanning the entire disk media for recognizable files. Each block of a computer storage media (*e.g,* hard disk drive) is searched for headers corresponding to known file types, such as a JPEG or Word document. Once a header is identified, a recovery utility extracts the corresponding data from the media until it reaches what it perceives to be the end of the file. This method is often referred to as file carving. File carving can yield a large number of recovered documents, but it suffers from the following significant limitations—

- Only the content is recovered. The file name, creation/modification dates, and other metadata are not recovered
- Files are often split into multiple fragments, making it difficult to identify and assemble those fragments back into a complete file.

The second approach is to examine the file system's primary data structures to locate remnant records of deleted files within the file system that have not been erased or overwritten. In simple cases a file may be marked as deleted. In more complex file systems the remnant may have been left by the file system during regular maintenance and updating. Compared to file carving, this approach has the following advantages—

- Deleted files may be recovered more accurately and include associated metadata such as the file name, size, and date information
- Because the location of the deleted file fragments is often known, the starting and ending blocks do not have to be found *via* search/carving techniques
- Files that are split into multiple fragments can be concatenated to reconstruct the entire file.

The limitation of this approach is that remnant records may be difficult to locate or may be quickly over-written by the file system. Other challenges may also exist depending on the particular file system implementation.

A New Approach

Journaling of file system metadata is a relatively new feature of desktop file systems. A journal protects the file system from corruption when the file system is improperly unmounted, such as when the system crashes or an external drive is unplugged without first stopping the device. When the file system is mounted again, a stable state can be quickly established without requiring file system repair. To accomplish this, file system journaling performs the following—

1. Prior to making a change, the file system writes a duplicate copy of the metadata in the journal with related updates grouped in a transaction
2. When the file system is mounted, transactions that are included in the journal, but not on the disk, are replayed to the disk and transactions that are incomplete in the journal are discarded
3. The file system is now consistent without requiring a disk repair.

By design, the journal includes historical information about the file system along with current state information. Additionally, the journal contents are stored separate from the primary data structures and may contain

more remnant metadata than the file system's primary data structures.

Journaling file systems in common use include NTFS for Microsoft Windows NT, 2000, and XP operating systems, ReiserFS and ext3 for Linux variants, and HFS+ on Mac OS X. Deleted file recovery, including journal analysis, is a well-researched topic for Windows and Linux operating systems, however our research focuses on analysis of Mac OS X's HFS+ file system and specific issues related to it.

Tools for recovering deleted files on Mac OS X's HFS and HFS+ file systems have had limited success compared to recovery tools for other common file systems due to the HFS file storage method. HFS+ file records are stored in a B-tree data structure, which dictates that upon deletion a file record is immediately removed, rather than simply tagged as deleted.

Experiments and Results

Beginning with Mac OS X 10.2, Apple's implementation of HFS+ supports journaling of metadata. Our research investigated whether the journal contains historical copies of file records, some of which may correspond to deleted files, that could be used to accurately identify and recover deleted file data. Analysis of the journal confirmed that journal functions duplicate disk blocks in their entirety, and that directory records may be among them. Subsequently, we developed a technique for parsing the journal.

Mac Mini Volumes	
Volume 1	Boot volume on the internal hard drive
Volume 2	External Firewire drive
Volume 3	External Firewire drive
Volume 4	Mounted disk image

Table 1 Mac Mini Volume Information

MacBook Pro Volumes	
Volume 1	Boot volume on the internal hard drive
Volume 2	External USB drive, Time Machine backup drive

Table 2 MacBook Pro Volume Information

Our implementation can be summarized as follows—

1. Scan the journal file chronologically, starting from the oldest point
2. Identify Catalog file nodes by examining headers and record pointers
3. When a Catalog node is found, examine each record and locate it in the active file system Catalog. If the record cannot be located, conclude that it is deleted
4. When a deleted file record is found, check the file system's block allocation table to see if the file's blocks have been reused by another file. This determines if the content data is still available for recovery.

Experiments were conducted with two test systems to gauge the following—

- Length of time the data persists in the journal file
- Efficacy of recovering files in real-world scenarios.

The test platform consisted of a Mac mini and a MacBook Pro, described in Tables 1 and 2 respectively.

The systems were monitored for 8 hours, during which typical user tasks were performed including web browsing, reading email, and editing documents. The Mail application on the MacBook Pro was configured with 5 email accounts, it checked email every 5 minutes, and it received and filed approximately 200 emails during the monitored period.

The first area of focus is the length of time the data persists in the journal file. In general, the boot volume of a system is subject to regular user and system activity, so the time window is relatively short. On the Mac mini's boot volume, the window was typically 5–10 minutes. On the MacBook Pro, the window was typically about 30 minutes, which was due to a larger hard drive and a corresponding larger journal file. The other volumes on the Mac mini saw limited activity and the time window was greater than the 8-hour period

of testing. The Time Machine volume of the MacBook Pro was idle between backups, but when a backup was in progress, the time window was a short 30–60 seconds.

The second area of focus was on actual recoverability of deleted files. Using a prototype implementation of the technique, recovery was attempted at the end of the 8 hour period. The experiment resulted in the recovery of a number of files across all volumes as shown in Table 3.

Recoverable Files Per Volume		
MacBook Pro	Volume 1—Boot	67
	Volume 2—Time Machine	3
Mac Mini	Volume 1—Boot	14
	Volume 2—External	119
	Volume 3—External	36
	Volume 4—Disk Image	162

Table 3 Recoverability of Deleted Files

Limitations of the Approach

While our technique did result in the recovery of recently deleted files, it has the following limitations—

- The journal represents a relatively short window of time, typically between a few minutes of file system activity and several hours
- Deleted entries are not guaranteed to be in the journal
- The full path to the file may not be retrievable. The path is constructed by following the parent folder repeatedly until the root of the file system is reached. If any parent has been deleted, the path will end at that point
- The file deletion time is not known. Files are determined to be deleted by deduction, so there isn't an explicit event in the journal to identify the time at which it was deleted
- There is no guarantee that the allocation blocks contain the same data as when the file was active. It is practical to check if the file's blocks are currently in use, but it can't be determined if the blocks were reused then subsequently freed after the file was deleted
- The technique is not applicable if the file system is not intact. Because it relies on searching the active file system, a corrupted or deleted file system will cause searches to fail.

Conclusion and Future Work

Despite the limitations, the deleted file recovery technique represents an innovative method to advance the state of the art of Macintosh computer forensics analysis and digital evidence collection. Future research will work to overcome the limitations of this technique and provide accuracy measures to further assist in investigating computer-aided criminal activity, network intrusions, and computer network attacks. ■

About the Authors

Aaron Burghardt | is a software engineer engaged in research and development of forensic solutions on behalf of Federal clients. His work focuses on Macintosh and iPhone technologies.

Adam Feldman | has 23 years of professional experience in the areas of information and computer security, software engineering, and investigative technologies—including digital data and text analysis and computer forensics. He is currently providing program management, technical and strategic direction, and thought leadership for several text analytics/mining initiatives and computer and network forensics projects for federal government clients.

Letter to the Editor

Q ***I recently read a fascinating article in a previous edition of your* IAnewsletter *and had some questions about the topic. How would I go about having my questions addressed?***

A Thank you for your inquiry. We, in IATAC, do our best to solicit and select articles with topics that will be of interest to the greater IA community. We also expect that some of these topics will generate questions, comments, and even concerns. In order for us to address these, we simply ask that you contact us either by email at *iatac@bah.com* or by phone to 703/984-0775. Depending on the specific topic and inquiry, either someone within IATAC will respond or we will work with the author of the article for answers. Either way, trust that we will do our best to respond to any question, comments, or concerns that may arise from one of our *IAnewsletter* articles. ■

Data Protection Life Cycle, Part 1—Knowing What You Have

by Allan Carey

Data protection is more of a cultural, business, and legal issue, than a technological problem. While this statement holds true time and time again, most organizations attempt to attack the problem in reverse. Protecting data starts with understanding its life cycle, classifying it, and identifying where it resides and who is accessing it. For the purposes of this discussion, data refers to data residing on unclassified networks. Once data is classified, various techniques can minimize its loss, including use of data loss prevention (DLP) products, which are good at stopping ill-advised user activities, but not targeted malicious attacks; and database security controls.

Before an organization begins a data classification project, it should consider its data life cycle. One example of an organizational data life cycle is data creation, aggregation, daily use, archiving, and destruction. The data life cycle will be different for each organization from agency to agency, agency to enterprise, and enterprise to enterprise. Factors that influence the data life cycle include complexity of business processes and types of data the organization generates. As an organization develops an enterprise data life cycle, the following are key considerations—

- **Risk**—At each phase in the life cycle, there are different levels of risk. This should be recognized and the risk levels should be documented.
- **Data owners**—It is important to identify a data owner for each phase. There may be one owner for each phase, or multiple owners.
- **Databases**—A data life cycle can also apply to databases; *e.g.,* creation of a database, data processing/online use, server offline situations, in transit, and backup.

There are numerous challenges, however, in successfully implementing data classification.

- **Senior management support**—Critical to success are obtaining management buy-in and having a champion who can drive the initiative.
- **The location of data**—While it is easy to develop a data classification scheme, most companies don't know where their data resides. From a risk management perspective, it is important to identify where the critical data is stored. Institute for Applied Network Security (IANS) recommends using a business impact analysis (BIA) as a starting point if one exists.
- **Employee participation**—Although a data classification scheme may exist, employees are probably not adhering to the policy. Training and awareness can help address this problem.
- **Data in motion**—Labeling data with the appropriate classifications and then maintaining those labels, as data travels can be difficult.

Understanding what data are important is crucial to success. Each organization has its own definition of critical data. Information security professionals should work with the business owners to identify the high-importance and high-risk data. IANS recommends no more than four levels of classification because our research has shown two is too little and anything more than four becomes confusing to users.

- Ask business owners the worst thing that could happen if certain data left the organization or became lost or stolen
- Perform a process flow analysis to determine the worst leaks that can be found
- By conducting research, collecting data, and developing a DLP plan, the information security team can demonstrate to business leaders how they are exposed through data loss and how much it will cost to reduce this risk.

Before implementing any sort of DLP program, engage Legal advisors. Keep in mind that from a legal and regulatory perspective, any critical data that is

▷▷ *continued on page 30*

The EPOCHS Project:

Creating a Framework for Managing Future Electric Power Systems in a Secure, Efficient, and Reliable Manner

by Kenneth Hopkinson and Stuart Kurkowski

Introduction

Stress on the electric power grid continues to rise in the current deregulated environment. Demand for power has grown both with the increasing population and the greater electrification of the economy. Despite this, the transmission capacity of the grid has remained largely static, meaning instabilities and disturbances in the future will have significant impacts and will increase in frequency of occurrence. Currently, the lack of information within regions of the electric power grid serves as a real limitation to detecting and effectively reacting to system instability. The difficulty level rises by an order of magnitude when attempting to react to problems that originate outside of a region that might be affected by them. The inadequacy of the current system for monitoring the power grid was highlighted in the blackout that occurred in August 2003. [1] The Midwest ISO's (MISO) state estimation system stopped its updates when SCADA information failed to arrive on lines that failed in electric utility CINergy's domain. The 2 power system operators at MISO failed to notice the resulting alarm and the new SCADA information was not being displayed. The monitoring system failure was a serious one, and was one of the major contributing factors of the blackout within the ISO's region. It is important to note, though, that none of the neighboring regions have a clear picture of the state of the Midwest ISO even under the best conditions. This lack of shared state information facilitated the spread of the blackout far beyond Ohio-based First Energy's borders. There are many signs that the power industry is turning towards next generation communication systems in order to meet the increased demands that are being placed on the electric power grid and region to region communications. Recent standards such as the IEEE Utility Communication Architecture (UCA), IEC 61850, and research efforts such as the use of Wide Area Measurement Systems (WAMS) in the Western United States are just three major examples of the active interest in the electric power industry. These standards point towards the future adoption of a private Utility Intranet based on Internet technology to improve the efficiency and reliability of the power grid. The Utility Intranet is likely the effort to improve upon the monitoring, protection, and control of individual utilities and, with communication standards such as the previously mentioned IEC 61850, will lead to the interconnection of the utilities' data networks in the same way that the electric power grid has become integrated over time. The introduction of a Utility Intranet has many potential benefits such as increased information sharing, greater protection and control of the grid, and the enhanced ability to share power in complex situations including bilateral load following. However, great care must be taken to ensure that network capacities, communication protocols, security, and Quality of Service (QoS) requirements are appropriately managed to ensure that the Utility Intranet can meet the demands that are placed upon it. Most of the installed base of protection and control equipment in the grid has minimal security embedded within it and a low tolerance for communication delays, which makes it difficult to strengthen using firewalls, intrusion detection devices, and other security mechanisms. There are also difficulties in making effective use of the network itself.

It is natural to assume that electric power protection and control systems with more data, operating over faster communications networks, will be more effective than their predecessors. Yet, the Internet was not designed for safety- and time-critical applications. Layering the needed mechanisms over Internet protocols such as TCP/IP is likely to be an extremely challenging undertaking. Investigations into the characteristics of a future Utility Intranet are required to understand its implications before far-reaching decisions are made.

Trust Platform

SCADA systems used to manage complex utility networks, often with thousands of monitored nodes, have to be capable of reliable and accurate real-time or near real-time responses to fluctuations and emergency situations. Traditionally, each company had its own proprietary systems and protocols from various vendors with no community standards. Interoperability and security often took a back seat to efficiency and functionality. Many companies felt secure due to the uniqueness and complexity of their systems. Now deregulation has broken up many of the previously held monopolies so that each privately-owned company specializes in only one function (*i.e.* generation, transmission, or distribution). It has also served to increase competition resulting in a greater need for management efficiencies and the protection of company-sensitive data. So, in recent years, utilities have begun to move from the 3 proprietary systems and protocols that once dominated the industry toward open, networked communication standards for control and data acquisition, patterned after the efficiencies and lower cost of Internet technologies.

Often power engineers with a desire to maintain finely-honed processes and operational requirements raise concern that the majority of information technology (IT) security mechanisms used in networks, like the Internet, will upset the delicate balance in SCADA networks. IT personnel familiar with security mechanisms used to defend office networks, such as firewalls, intrusion detection devices, and encrypted traffic standards, see them as the most secure measures for protecting systems against threats such as malicious code and online exploits, but they are mostly delay-tolerant networks. Thus, both parties are at odds regarding the role, priority, and implementation of security countermeasures.

The focus of the Trust System research has been to investigate the claims from both sides with respect to the feasibility of employing common, network security mechanisms to real-time SCADA and near real-time wide area measurement systems. Utility Intranets could be applied to the power system in the United States and Canada, across interconnected grids in the European Union, and in other major regions across the world. Utility Intranets will allow shared information about the state of different connected systems to facilitate enhanced protection and control through greater knowledge and coordination between regional control authorities. While there are many perceived benefits of this shared information, security concerns will necessarily rise as the grids' information systems become interconnected. There will also be the potential for shared protection and control between utilities, which raises even greater security concerns.

It is assumed that future Utility Intranet SCADA networks will resemble modern IT network architectures. This collaborative Trust System, shown in Figure 1, is a hybrid solution comprised of the leading IT security mechanisms and standard IP protocols while focusing on the distinct requirements of the SCADA community, such as the need to allow increased cooperation and information sharing in protection and control systems without disrupting the critical operation of these systems.

Early experiments have been run, based on published performance figures, in order to illustrate the operation of the Trust System in a sample scenario. The messages defined for use in this research contain the additional overhead of TCP/IP, larger IPV6 addresses, and encryption. In this way, the research results accurately represent the delay for Trust System evaluation of real-world messages of the same general size.

This early research has shown that, even with the overhead of TCP/IP and UDP/IP communications, Internet Protocol Security (IPSec) encryption, firewall rules, format check, and access control functions, the recommended security schema of the trust system can perform within near realtime and at the high end of real-time response time constraints as long as they are carefully placed with those constraints in mind. It is deduced that with further optimizations, the schema can be improved to perform satisfactorily for real-time SCADA systems.

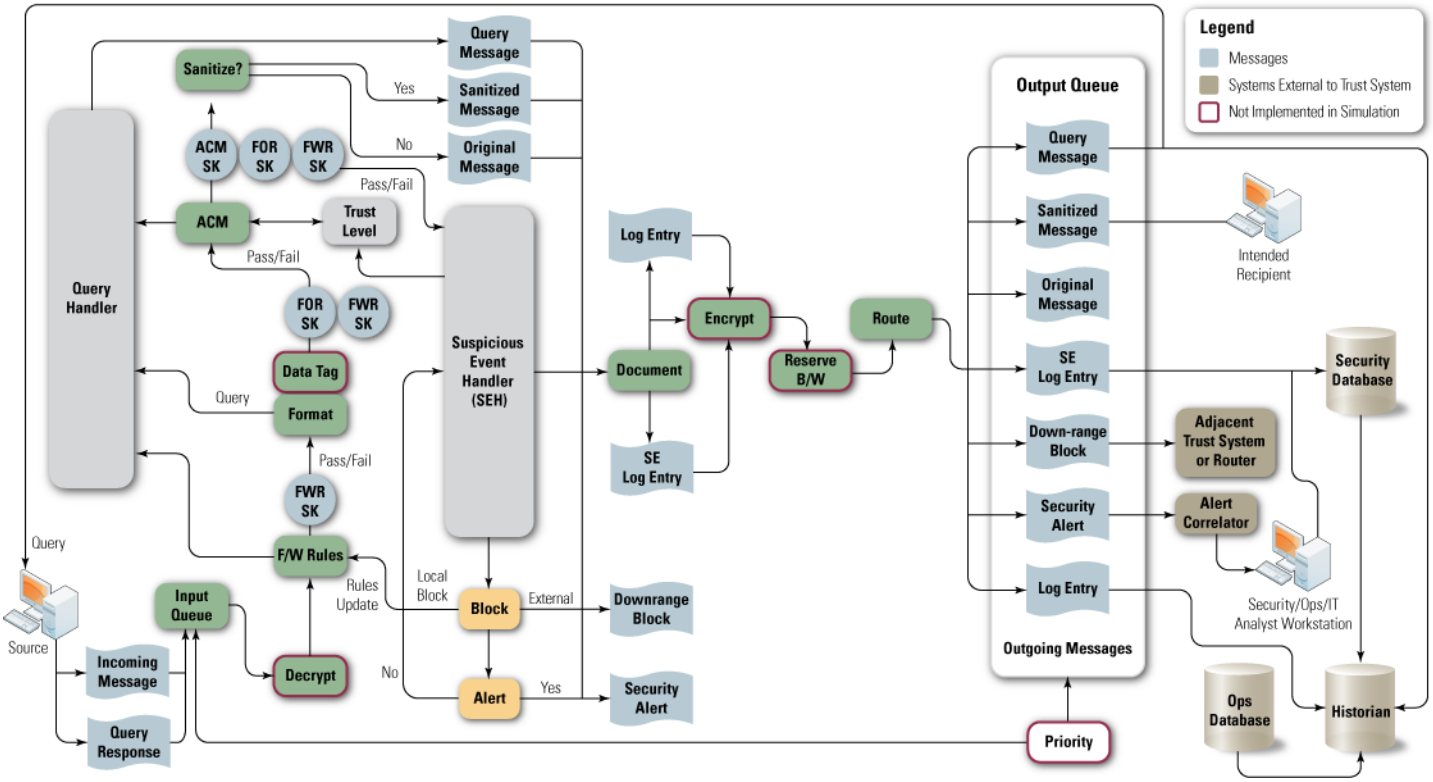


Figure 1 The Trust System [10]

The EPOCHS Platform

The restructuring of the electric power system, the creation of competitive markets, and the introduction of new regulatory mechanisms are now well-established trends. Understanding how the electric power grid operates, and in particular how to monitor and control it, are necessary preconditions to achieving reliability and fairness in the presence of these changes. Yet, electric power simulators today do not model the network communication patterns seen in modern protection and control systems. At AFIT, in conjunction with earlier work conducted at Cornell University, we have developed a platform to allow us to investigate utility network traffic models, security standards, and communication protocols and their reaction to simulated power systems. The electric power and communication synchronizing simulator (EPOCHS) federates, or combines, GE's PSLF [2] and Siemen's PSS/E [3] electromechanical transient simulators, HVDC Manitoba's PSCAD/EMTDC electromagnetic transient simulator, [4] and the University of California at Berkeley's network simulator 2 (ns2) [5] together to allow users to study electric power protection and control systems that depend on network communication. [6] Using EPOCHS, electric power engineers can look at protection and control systems in the presence of electromechanical transient situations that depend on communication, using PSS/E in conjunction with ns2, or electromagnetic transient situations, using PSCAD/EMTDC in conjunction with ns2. The individual simulators are seamlessly integrated from a modeler's perspective. Users are able to create software agents to mimic the behavior of real systems. Using agents, modelers are able to interact with an integrated world containing both electric power and networking state, can take local measurements, set electrical state, and can send or receive messages across a communication network.

Agents can also model security technology, such as the Trust System

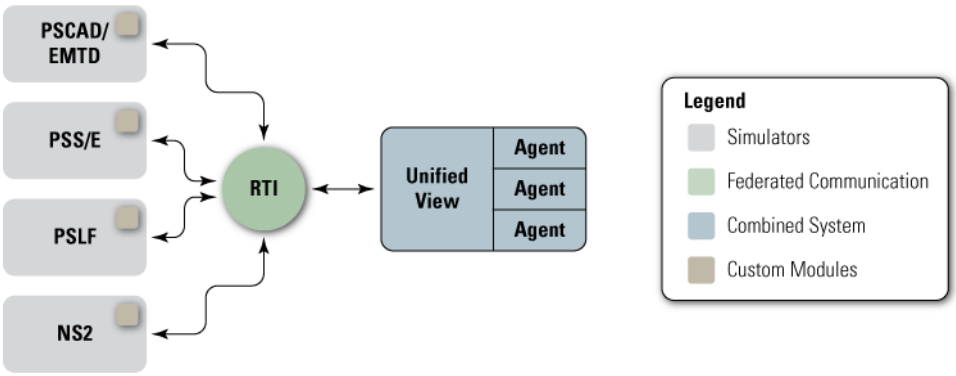


Figure 2 The EPOCHS Simulation System

outlined earlier, to give an accurate view of how these security mechanisms impact a working electric power system. A module known as the run-time infrastructure (RTI) ensures that the simulators are properly synchronized so that if an event happens at simulation time t in the power or network simulator then it occurs at the same time in the other simulator. The combination is a powerful one and allows studies into the effect that communication has on the operation of protection and control systems that are dependent upon it. Figure 2 gives an overview of the EPOCHS simulation system.

Using EPOCHS, experiments have been performed to show how a Utility Intranet could be used to allow network communication to enhance the capabilities of protection and control systems. These experiments have been performed in areas involving special protection systems used to prevent voltage collapse, [6] zone 3 backup protection relays, which can coordinate with one another to prevent relay misoperations, [6] and in grid monitoring systems designed to provide shared information between regions of the electric power grid. [7] These experiments all show promise for the future use of protection, control, and grid monitoring systems augmented with network communication.

AFIT, using EPOCHS, is looking at the implications of protection and control systems which have dependencies on network communication. The special protection study mentioned in the previous paragraph briefly looked at the implications if network packets were lost due to network congestion. Another study looked at the impact of packet losses on the operation of zone 3 backup protection systems that depend on network communication. [9] The latter two studies are notable because they begin to show the impact that protocol selection and network congestion have on the performance of protection and control systems that depend on network communication. The EPOCHS platform also has the ability to show the benefits, and drawbacks, of security platforms, like the Trust System, on realistic protection and control systems, and we plan to exploit this capability as our work progresses.

Background Traffic and Quality of Service in Utility Networks

The Utility Intranet, outlined earlier, is likely to be employed for many different purposes by the electric power community. Some of the most likely uses of the Utility Intranet are outlined in the eight categories shown in Table 1.

The "Distribution," "Packet Size," and "Rate" columns in Table 1 show a model, based on the traffic types which can be used in simulations in order to demonstrate the effects of background traffic on the operation of electric power protection and control systems. While the traffic over the future utility intranet may differ from that presented in Table 1, it has value both as a first attempt at quantifying the type and form of traffic that is likely to appear and in the ability to use the model to demonstrate the impact of inherently low-priority traffic when critical communication traffic is present in the network. These models were used in key experiments to demonstrate the impact that the management and security of the utility networks can have on the performance of critical protection and control systems. These models also serve as a stepping stone towards further work where the realistic impact of security environments, like the Trust System, can be measured with the aid of the EPOCHS platform.

Background Traffic Type	Distribution	Packet Size	Rate
SCADA	Constant	64 Bytes	1 every Second per Bus
Power Quality Data	Poisson	35 Bytes	1 every Second per Bus
UCA 2.0	Poisson	128 Bytes	1 every 20 Seconds per Bus
Power Trading	Constant	1,400 Bytes	1 every 2.2 Seconds per Bus
Internal Comm	Poisson	1 Mbytes	1 every 0.2 Seconds per Bus
Office–Substation	Poisson	64 Bytes	1 every 10 Seconds per Bus
Event Notification	Poisson	2.4 Mbytes	1 every 10 Seconds (Bus chosen at random)

Table 1 The Background Traffic Rates for Low Traffic, consisting of white sources, Medium Traffic, consisting of both white and light gray traffic, and Heavy Traffic, which consists of white, light gray, and dark gray traffic sources

Visualization Work

At AFIT, we have developed an integrated framework that mediates the viewing of ns2 network traffic from the simulation engine to a graphical display in lock-step with the simulation execution. [11] This work renders the graphical representation of the network entities, their location, characteristics, links, and connections, as well as packet flows in the network as ns2 generates the activity. We have a visualization mediator that extracts activity information from ns2 during execution and an RTI component that synchronizes the event steps across various simulators.

The robust framework, shown in Figure 3, has been tested against 16 different trace formats from ns2 and many visualizations. The framework can render physical layout plots, locating nodes in their geographical locations showing physical relationships between nodes. Additionally, networks operate in a virtual world of connections, priorities, and speeds. The framework can also render the network using force-directed graphs which move network entries into groupings of differentially weighted springs attracting nodes of a certain characteristics and

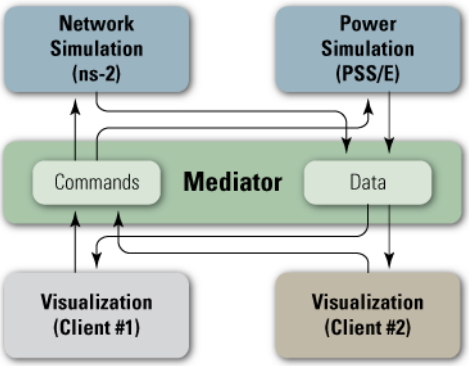


Figure 3 Visualization Framework Architecture

repelling others. This force-directed graph layout allows analysts to monitor changes in the network instantly. It even allows the indication of events prior to their full onset state. For example, if a network link is reaching capacity and bandwidth is the current characteristics used for the force-directed graphs, the nodes will continue to move together approaching 100% capacity, prior to repelling each other at 0% capacity when the link breaks. These changes cause dramatic topology differences, easily viewable by the analyst.

The framework links the visualization engine with ns2 in both directions. The mediator allows information from ns2 to be displayed on screen, as described in the previous paragraphs, but also allows interactive widgets to appear on screen for information to be passed back in to the simulator. When combined with the EPOCHS platform, the system gives an interactive framework to allow users to control power communication systems interactively. Power communication systems have traditionally been relatively simple, which has made such control systems unnecessary. As network traffic management, security, and protection and control traffic increase in complexity, we expect that such systems will become a key part of utility control centers. The system also allows explicit control and feedback of security components and critical factors such as the relative trustworthiness of SCADA and other network components throughout the system so that operators can factor such considerations into their decision-making.

Conclusion

The work performed as part of the EPOCHS project has only begun. There are many aspects to the ongoing work at AFIT that we plan to begin and/or continue. We see the project contributing in many ways to the public good in terms of adding critical infrastructure for research and education, and in added societal benefits. The summary of these benefits are as follows—

Infrastructure for Research and Education

Our on-going work will strengthen power grid companies and university partnerships by addressing important research questions that are of mutual interest. These relationships serve to broaden student perspectives by allowing them to gain exposure to practical problems arising in real contexts, and to benefit from the expertise of practicing scientist and engineers.

Societal Benefits

The on-going work will enhance and accelerate the design, analysis, control, and implementation of Internet-based power grid control systems that are becoming embedded in our society. A better understanding of the security implications of this transition will result in a better safer national power grid system. More broadly, the intelligent use of the power grid control system in a faster, safer manner, will lead to improved national security, economic stability, provide early warning of issues, and reduced lost of life for the national as a whole. ■

References

1. US-Canada Power System Outage Task Force, "Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations," April 5, 2004.
2. General Electric, "PSLF Manual." vol. 2003, 2003.
3. Shaw Power Technologies Inc., PSS/E 30 User's Manual. Schenectady, NY, USA, 2004.
4. Manitoba HVDC Research Centre, PSCAD/EMTDC Manual Getting Started. Winnipeg, Manitoba, Canada, 1998.
5. L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidermann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu, "Advances in Network Simulation," IEEE Computer, vol. 33, pp. 59–67, May 2000.
6. K. Hopkinson, X. Wang, R. Giovanini, J. Thorp, K. Birman, and D. Coury, "EPOCHS: A Platform for Agent-based Electric Power and Communication Simulation Built from Commercial Off-The-Shelf Components," IEEE Transactions on Power Systems, vol. 21, pp. 548–558, May 2006.
7. P. Birman, J. Chen, K. M. Hopkinson, R. J. Thomas, J. S. Thorp, R. Van Renesse, and W. Vogels, "Overcoming Communications Challenges in Software for Monitoring and Controlling Power Systems," Proceedings of the IEEE, vol. 93, pp. 1028–1041, May 2005.
8. K. Hopkinson, G. Roberts, X. Wang, and J. Thorp, "Quality of Service Considerations in Utility Communication Networks," IEEE Transactions on Power Delivery, To Appear.
9. R. Giovanini, K. Hopkinson, D. V. Coury, and J. Thorp, "A Primary and Backup Cooperative Protection System Based on Wide Area Agents," IEEE Transactions on Power Delivery, Accepted for Publication.
10. G. M. Coates, K. M. Hopkinson, S. R. Graham, and S. H. Kurkowski, "Collaborative, Trust-Based Security Mechanisms for a Regional Utility Intranet," IEEE Transactions on Power Systems, vol. 23, pp. 831–844, August 2008.
11. J. M. Belue, S. H. Kurkowski, S. R. Graham, K. M. Hopkinson, and R. W. Thomas, "Research and Analysis of Simulation-Based Networks through Multi-Objective Visualization," in Winter Simulation Conference, Miami, FL, USA, 2008, p. To Appear.

About the Authors

Kenneth M. Hopkinson | received his PhD in Computer Science from Cornell University in 2004. He is currently an Assistant Professor of Computer Science at AFIT. His research interests include fault-tolerant and distributed systems, networking, and simulation. His email address is *kenneth.hopkinson@afit.edu.*

Stuart H. Kurkowski | received his PhD in Mathematical and Computer Science from Colorado School of Mines in 2006. He is currently an Assistant Professor of Computer Science at AFIT specializing in software engineering, visualization, and simulation. His email address is *stuart.kurkowski@afit.edu.*

Cyber Defense Branch Takes Part in NSF Workshop in Beijing

by Kevin A. Kwiat

Beijing, China became the venue for the first Workshop on Cyber-Physical Systems (WCPS) in June 2008. Sponsored by the US National Science Foundation (NSF), WCPS attracted leading researchers from all over the world who were coincidentally in Beijing attending the heavily cyber-related International Conference on Distributed Computing Systems (ICDCS). Co-locating WCPS with ICDCS allowed the newly-founded workshop to leverage off the 27-year history of the conference. Below is the workshop logo—

Technically, cyber-physical systems have emerged from distributed computing systems. Distributed computing systems are the interconnection of computers over a network creating a shared processing of a task. From the coupling of many, yet geographically dispersed, computers comes the notion of cyber-physical systems (CPS)—the integration of information systems with their physical environment. Cyber-physical systems are distributed computing systems but with some unique properties—they are physical and engineered systems whose operations are integrated, monitored, and controlled by a computational core. The science of this field is moving rapidly, in large part because of the proliferation of affordable network technologies and the commercial markets that employ them in order to bring more classes of physical processes under computer control. Conversely, through direct feedback, the physical processes also impact the computers. The NSF frames CPS as systems that will transform how we interact with the physical world just like the Internet transformed how we interact with one another.

From the coupling of many, yet geographically dispersed, computers comes the notion of cyber-physical systems (CPS)—the integration of information systems with their physical environment.

Internet security concerns are compounded in CPS because exploitation of a CPS vulnerability can have immediate and dire physical consequences on an unprecedented scale. Responsively, the majority of papers presented at WCPS dealt with security; yet modeling of CPS received the second-most representation paper-wise. Researchers contend that modeling must capture the multi-disciplinary nature of CPS. Specifically, within CPS the "C" people (comprised of computer scientists and engineers) and the "P" people (as represented by physicists, mechanical and civil engineers) need a common language that automatically translates and integrates disparate technical treatments of a CPS.

The cyber-relevance of the workshop warranted participation by the Air Force Research Laboratory Information Directorate's (AFRL/RI's) Cyber Defense Branch. Prior to participation, however, the particular venue of the workshop meant involvement by several offices to obtain the requisite approval including the Air Force Office of Scientific Research's International Office, Secretary of the Air Force International Affairs, and Office of the Under Secretary of Defense for Policy. Dr. Kevin Kwiat of AFRL/RI's Cyber Defense Branch represented the Directorate. His keynote address at the

▷▷ *continued on page 30*

Incorporating Flow-Based Behavioral Analysis Inside Agency Networks

by Frank Doane

Abstract

The US Government is embarking on an ambitious endeavor that will substantially change the way government networks communicate with the outside world and how they are protected from external threats. The reduction of Internet gateways through the Trusted Internet Connection (TIC) initiative, and the standardization of threat management and incident response through development of core requirements for Trusted Internet Connection Access Provider (TICAP) capabilities and the management of in-cloud security by the United States Computer Security Readiness Team (US-CERT) under the EINSTEIN program, ushers in a new era of unprecedented inter-agency collaboration and network consolidation for our government's federated networks.

One of the keys to success for this effort is for information technology (IT) security professionals at all levels of the enterprise (within the CERT and TICAP down to the individual owners of systems) to be able to detect the presence of malicious users connecting to government systems, to communicate details about the nature of the attack and its source to the organizations responsible for carrying on further investigations of the event, and ultimately to be able to respond to bulletins notifying them of events that have had an impact on the networks they are responsible for protecting despite having missed the attack in the first place.

To meet this challenge in the current world where a zero-day threat is an everyday reality, agencies need an always-on technology that constantly monitors the activities of their own users and outsiders touching their networks. This can be used to detect anomalous and out-of-policy system activities to spot the presence of an attack early and make sure a record of each communication is preserved to support the investigation of events that are only apparent to the community protecting the network after the fact. Such capabilities can easily be added at all layers of the government enterprise by bringing flow-based network behavioral analysis into the technology architecture of the agency.

Flow works by leveraging the switching and routing infrastructure as a virtual surveillance grid. When remote flow collection technologies such as NetFlowTM and sFlow® are turned on, infrastructure components report records of the traffic they observe back to a collector for analysis. Network Behavioral Analysis systems are expert systems that process flow records into conversation-level log files that record the fact of a networked conversation and write the records to a flow table analogous to a telephone bill—showing who talked to whom, when, through which ports and protocols, and how much traffic passed during the conversation. These records are read over by an engine that builds a baseline of normal behavior for each system observed to have touched the monitored network. The system generates alarms on anomalous activities, keeps a record of all activities of each connecting host, and provides a graphical user interface to speed the time analysts require to isolate information about conversations relevant to an investigation.

Other components are utilized to provide the additional information

To meet this challenge in the current world where a zero-day threat is an everyday reality, agencies need an always-on technology that constantly monitors the activities of their own users and outsiders touching their networks.

necessary to speed investigation. This can include username and Media Access Control (MAC) address or hostname translation to IP address through integration with directory stores and DHCP servers; communication via *SNMP and syslog with other reporting systems to pull in additional data for correlation and gain more information about the network and hosts being monitored; and integration upstream to Security Information and Event Management Systems (SIEM) such as ArcSight or Enterprise Network Management (ENM) such as HP OpenView to leverage Behavioral Systems ability to compress flow records and isolate relevant communications to aid the use of data collected in a broader Service Oriented Architecture (SOA).*

This article surveys the capabilities of both the current tools used to protect government networks leveraging external monitoring programs (EINSTEIN and Centaur) and the capabilities of enterprise Network Behavioral Analysis (NBA) systems; and explores areas where use of a NBA, such as Lancope's StealthWatch, can enhance the ability of IT Security teams to detect suspicious behavior/zero-day threats and speed the incident response process in conjunction with EINSTEIN and Centaur to improve the Federal government's security posture under the emerging world of the National Strategy to Secure Cyberspace. The intended audience is anyone involved in operational security within the Federal government, to include US-CERT, Joint Task Force—Global Network Operations (JTF-GNO), TICAPs, Computer Network Defense Service Providers (CNDSPs), and agency-level network security and network operations personnel as well as those responsible for providing resources, technical guidance, and oversight to these communities.

What additional benefits would deploying NBA systems inside agency networks offer over and above the current capabilities for threat monitoring and response provided by EINSTEIN and Centaur?

The attacks we are discussing are primarily leveraging the involuntary recruitment of systems across the Internet *via* the infection of hosts who are invited or tricked into downloading exploit code—which in turn enables a controller to remotely access the owned system and delegate it to carry out tasks to accomplish the missions of the larger attack without detection. So the activities of both the coordinating controller and the infected systems will ultimately traverse the Internet for the success of the overall attack. Are there any advantages to watching inside the perimeter of the private network to find these attacks?

Yes. The reasons for adding NBA inside the agency networks as a defense-in-depth strategy are vital to the overall objectives of securing Federal government networks from current threats, and include—

1. Providing coverage for all communications traversing internal or private network space; eliminating missed attacks leveraging networked backdoors; and providing a comprehensive forensic record that shows not only the full extent of the compromise by picking up the presence of other recruits inside the network, but contains a record of any file transfers or other exfiltration events necessary to understand the impact of the event.
2. Earlier attack detection to leverage the predictability of host behavior as viewable only from the inside of the network to tune behavioral anomaly alarms that detect a single instance of internal host compromise and act as a "spotter scope" for teams in higher tiers conducting a comprehensive forensic analysis of widespread attacks. In addition to anomaly-based alarms, NBA systems offer customizable policy enforcement that acts as a trip wire to catch the presence of compromised systems by enabling agency analysts to define rules of the road for network hosts. These rules are often violated when a compromised system is remotely controlled by a party unfamiliar with established policies.
3. Increased speed of incident response by providing internal agency personnel responsible for

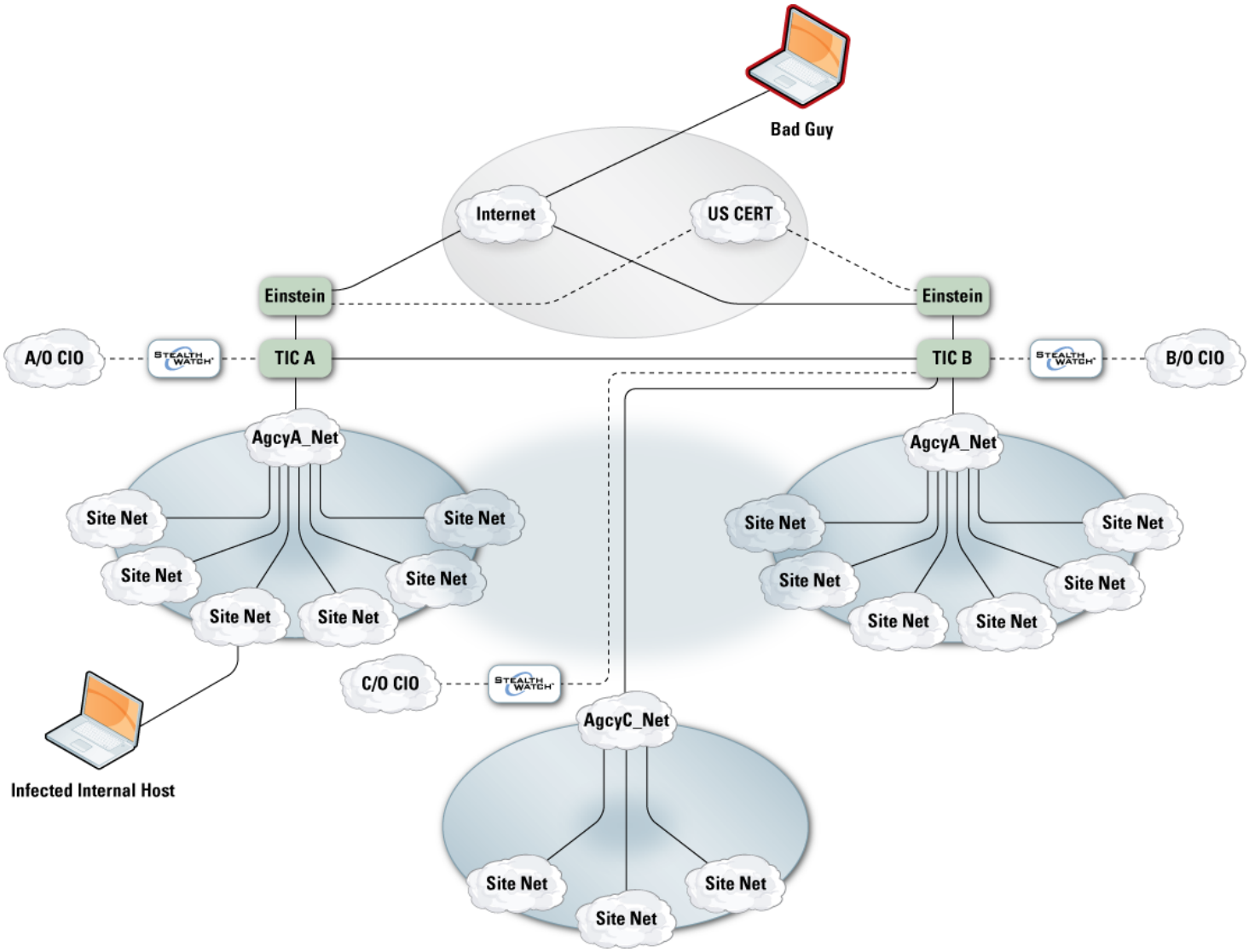


Figure 1

coordinating with the TICAP or CNDSP or the US-CERT or JTF GNO the ability to perform their own flow-based analysis of an attack touching inside the network to pinpoint the location and identity of the host actually responsible—a task that is impossible given the obscurity that NATing firewalls (firewalls performing Network Address Translation) and proxying devices create for higher-tier analysts.

Expanding Coverage Down to the User

The drawing below shows a high-level view of proposed coverage for the entire system with areas of responsibility broken out between agencies, TICAPs, and US-CERT. A similar model applies in the Department of Defense (DoD), with JTF-GNO, CNDSPs, and component agency networks acting at each level.

The drawing shows flow-based monitoring responsibility broken into two broad categories—above the TICAP inside the Internet, and below the TICAP inside government networks. Other monitoring systems provide visibility beneath the TICAP, and many mandated security controls ensure that systems attached to these networks meet a minimal standard for risk. What is missing when flow-based analysis is not applied at this lower level of infrastructure is visibility into host behaviors required to understand the normal operating conditions of the network—who connects to which systems how frequently, and the ability to leverage this knowledge to understand when a threat is present on an authorized system that has somehow managed to hide from the various systems protecting the hosts and networks on which they reside.

Adding flow-based monitoring at this level would provide a pervasive forensic trail to track any behavior of an internal system involved in a threat event and fully scope the compromise. It would also serve as an early warning system for threats that evade current controls, such as the walk-in threat where a system is compromised by physical access either through addition of a rogue system or the porting of malicious code through a removable drive; the mobile threat of laptops, personal digital assistants (PDA), or other systems that do not always reside on the protected network

and are thus susceptible to compromise while detached; and holes created by misconfigured systems such as networked backdoors, open wireless connections, bridging of network segments, rogue applications, or firewall ports left open.

Operationally, once flow is collected at this level the flow records are easily shared with analysts at higher levels. The key is to collect flow once and provide interfaces into the data to as many people as may require access, and establish the ability to limit access to only the data those users have reason to access. NBA systems, such as StealthWatch, accomplish this through role-based access permissions that limit analysts' access to data through functional roles defining what they can see of the data collected and data roles that control which hosts or network segments they have purview over. In addition, NBA systems include the ability to pull flow records from lower-level systems into aggregating systems such as SIEM or other reporting tools *via* a north-facing Simple Object Access Protocol (SOAP), Extensible Markup Language (XML)-based Application Programming Interface (API). This enables data to reside in the agency system until it is needed to support an investigation by a TICAP, CNDSP, US-CERT, or JTF-GNO, at which point only records relevant to the investigation would be pulled. Having this just-in-time pull capability provides the ability for systems to retain local storage of flow data until required for analysis, which mitigates some of the impact of bandwidth overhead required to support flow analysis throughout the infrastructure. It also mitigates privacy issues, and concerns over who owns the role of protecting the local network segments, the agency, or the service provider.

The benefits of flow to quickly pull together the facts of an event and isolate the individual hosts involved should be brought into agency networks and shared up to the TICAP analysts responsible for protecting these networks. Without this data, much of the work involved in investigating threats to validate that they are indeed events and extrapolate the full nature of the compromise involves arduous collection of data from the owners of various systems logging some part of the event—manually culling through logs and relying on guesswork for the parts of the picture left incomplete. NBA systems are ideal for this mission, as shown in the next section.

Integrating EINSTEIN and Centaur Outside Agency Enterprise Networks with Commercial NBA Inside to Turbo-Charge Government Detection and Response Capabilities

The Spotter Scope—Speeding the Time Needed to Detect the Existence of the Threat in the First Instance

NBA systems detect threats through deviations from normal behavior by networked hosts and through the detection of network policy violations. The "trip wire" policy enforcement capabilities of an NBA system monitoring internal networks allows agency security architects to design a set of "rules of the road" to which networked hosts must adhere when using agency resources. Violations of these rules will often flag the presence of a threat, since they often serve as a check of control systems such as firewalls to prevent known bad actions from occurring. For example, a firewall policy disallows communications of an internal host to a server farm asset across a certain port. A network administrator opens a bridged connection to the server farm subnet from the user subnet, and leaves the connection open after completing his work. The internal user gets "owned" through the exploits of a hacker whose reconnaissance of the network uncovers the backdoor. The hacker uses the backdoor to communicate with the server, bypassing the corporate firewall. The fact that a two-way communication is established between a server in the server farm and a host on the user subnet is alarmed against by the NBA system as violating established policy—and during the incident response, the hacker's presence on the network comes to the attention of security personnel.

Discovering this event *via* current controls would be difficult because the firewall never saw the traffic; the routing and switching infrastructure is not designed to report these events; the server did not realize the policy violation had occurred; and whatever permissions were used to gain access were probably accredited. Leveraging the flow-logs of the NBA system, the security team could then report this event to US-CERT or JTF-GNO, or to the TICAP or CNDSP responsible for monitoring their specific network. Once the hacker's actions have been validated at that level *via* a log file check for systems reaching external network systems, the full extent of the hacker's activities could be quickly assessed through queries into the flow files stored and managed by the CERT and JTF-GNO.

In the case of anomaly events, the key role of the NBA system is again to find the original malicious actor and leverage it as a starting point at higher levels to investigate the full scope of the attack. NBA systems are particularly good at using non-deterministic means of detecting threats. Instead of applying deterministic rules as in the policy enforcement model, anomaly detection is accomplished through the NBA system profiling every host observed to have touched the network—establishing thresholds for certain traffic characteristics that are tracked as potentially security-relevant (*i.e.,* numbers of syn packets sent without a response, numbers of concurrent sessions established, *etc.,* over periods of time) and then leveraging algorithms that look for patterns in behavior indicative of a network attack or compromise by observing traffic that exhibits these characteristics once the threshold established as "normal" is exceeded for each host. The key to anomaly detection is its "fuzziness" or ability to point at something that appears odd, as opposed to signature-based detection that relies on deterministic coding of rules meant to detect threat events in progress. In the modern network threat arena, the ability to catch the actual attack has been greatly diminished by the use of "pull threats" such as pfishing, redirecting traffic to websites through DNS exploits, and other means of

drawing or tricking users to visit a site where they unwittingly load exploit code to their systems. This is in contrast to the old model, where the hacker sought out victims through scanning, then "pushed" the exploit to vulnerable hosts.

Anomaly events look at general categories of behavior consistent with threats across the network, such as scanning, spamming, Distributed Denial of Service (DDoS) and high-volume Domain Name System (DNS) queries. More often than not, these systems are picking up not the actual attacker, but one of its minions of "bots"—infected end systems tasked to participate in the attack by a controller who earlier had malicious code injected into the system through pfishing or other recruitment techniques. Systems such as Lancope's StealthWatch have further reduced the simplicity of detecting the presence of compromised hosts by looking at the aggregate of bad behavior emanating from a particular host without reference to a pattern of attack through index-based alarms that point to net bad actors or net targets of attack.

Once alerted to the presence of a bad actor, an NBA system provides a forensic record of all activities of the host over time whether or not suspicious events are associated with the particular communications involved. It is easy to track the source of infection back to a particular host on the outside merely by expanding the search of flow records back in time and examining suspicious "calls out" to external systems (*i.e.,* communications across IRC ports or at suspicious times of the day).

Once detected, the presence of a bad actor can be quickly validated and fully examined under the microscope of the larger flow-monitoring technologies of the Network Situational Awareness (NetSA) tool suite by US-CERT analysts or JTF-GNO. Having a starting point to launch an investigation and quickly pull together all points in the government impacted by the event is of critical benefit to the analysts at this level. While global event correlation systems such as the NetSA tools can detect threats by seeing patterns emerge on a global scale and "connecting the dots" of multiple infected systems back to a controller, it helps them enormously to know where to start.

Marrying multiple complementary detection technologies produces greater context around a security incident. For example, a global correlation system will identify the presence of a botnet across a system when multiple hosts from multiple networks—the bot army—are exhibiting similar or dissimilar behaviors associated with threat events, DNS or scanning-based reconnaissance, SPAM production, high volumes of connections to other systems, *etc.,* across the network but are periodically communicating with a single host that they all share in common—the bot controller. This behavior is only viewable once the monitored space is enlarged to gather enough data points to see a certain number of bots. Conversely, an NBA system monitoring component enterprise networks within the system will be able to detect the instance of a single bot engaged in suspicious behavior such as scanning outside the enterprise network, doing DNS queries, sending a high volume of email traffic, talking on IRC channels, *etc*. Having different systems reporting both the big-picture view of the bot army as the whole and the smaller picture of the single infected system in many different instances increases the likelihood that an attack will be caught earlier and minimizes the possibility of the attack escaping detection entirely.

The Hunting Dog—Speed Time of Incident Response

After an event has been validated by US-CERT/JTF-GNO, the next step is generally to announce its presence to the community of security professionals responsible for monitoring agency networks. This has taken the form of bulletins specifying the location of the controller (including URL and/or IP address), details about how the attack was accomplished, and details about the agency systems involved. Once the agency security personnel receive this bulletin, a new hunt begins. The easiest part is to reconfigure existing control systems such as proxies or firewalls to prevent further contact with the bad external actor. Discovering which internal end-systems are impacted, and isolating them on the network in the present, becomes an arduous task that involves the collection of data owned by multiple internal network professionals and reading through this data to understand where the impacted systems are today to affect remediation actions.

Looking at how this could be accomplished with an NBA system will show how much easier the incident response portion of the investigation can be made when NBA is brought into an agency to analyze internally collected flow. First, a policy enforcement rule can be established on the internal NBA system that looks for attempts to connect out to the controller to pick up systems that were compromised but had not previously reached out. This rule can be established to look for connection attempts instead of completed connections, since the connection will be presumably blocked at the firewall. Second, a quick flow analysis of the flow records for all internal systems can be run to look for any past connections to the malicious actor outside. When that list is returned, it can be further refined by looking at the URL to make sure the hosts involved were communicating with the website involved instead of another website residing on the same IP address, since many are in hosted environments. Once the communication has been validated, the internal system must be isolated. The quickest way to accomplish this is to use the NBA system's ability to integrate with directory stores and Dynamic Host Configuration Protocol (DHCP) servers to look up the username, hostname, or MAC address associated with the IP address of all internal systems observed to have contacted the external host at whatever point in time their contact occurred. From there, the analyst needs only look up the current IP address for the same username or hostname/MAC address

to identify its current location on the network. NBA systems will often include information about infrastructure devices connecting the user to the network so the actual switch or router and interface information can be isolated given the IP address, and the user can be quickly taken off the network through reconfiguration of the infrastructure devices.

Anyone having experience with running down the information required by a CNDSP, JTF-GNO, or US-CERT today will understand how much time and effort is saved through the process outlined above. The current alternative involves hours or days of collecting log data and manually culling through it for answers, which often leaves the job half-completed. ■

About the Author

Frank Doane | is the Federal Sales Representative for Lancope, Inc. Mr. Doane is an attorney in the state of Virginia, a graduate *cum laude* of George Mason University School of Law and holds a BA in History from Knox College in Galesburg, IL.

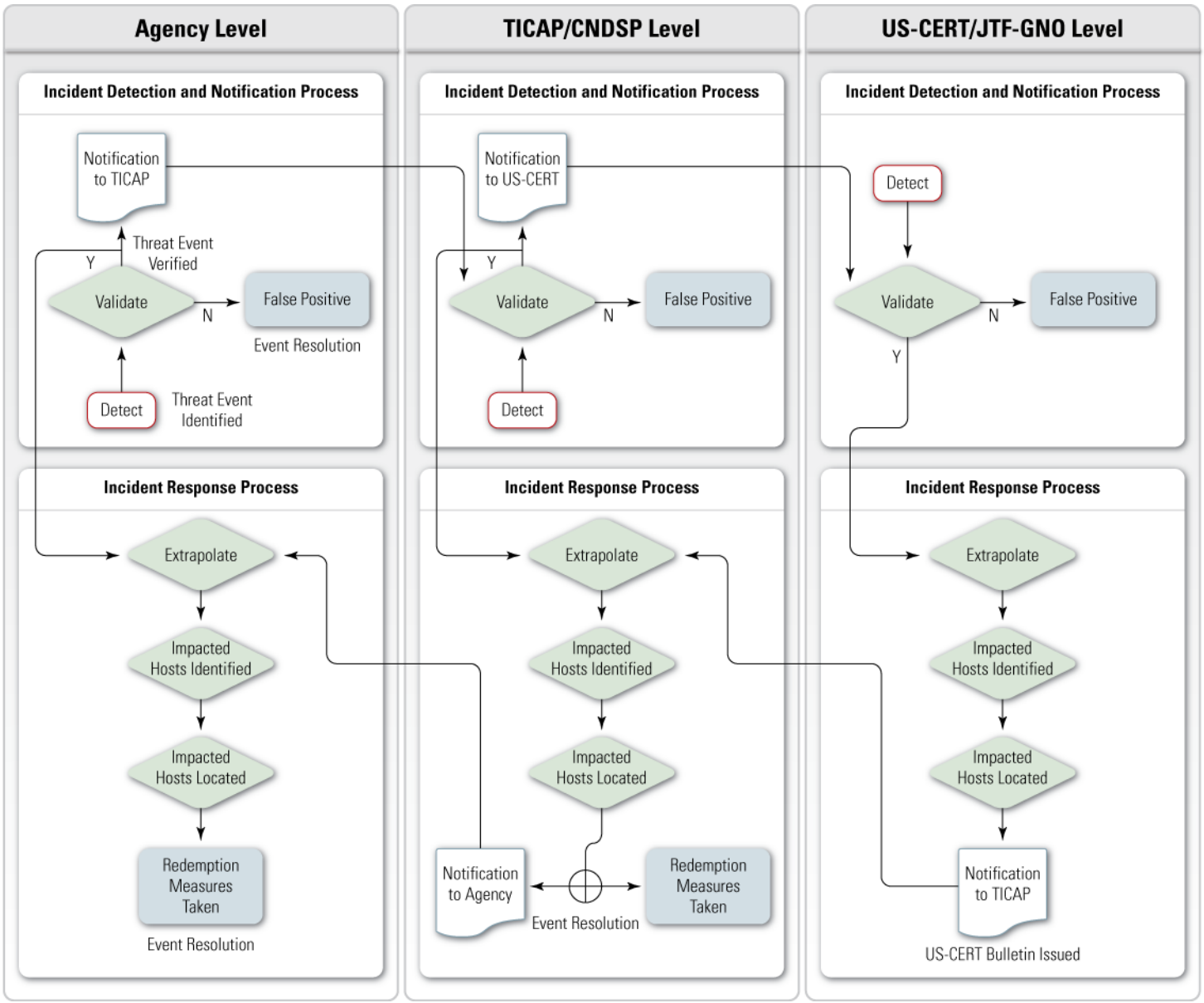


Figure 2 Putting it all together—proposed model for integration between US-CERT/JTF-GNO – TICAP/CNDSP and downstream agency security professionals leveraging NBA on agency networks

▷ continued from page 17

important to the business (*e.g.* such as source code or a key word), but is not regulated and does not have legal implications. Before trying a DLP solution, put compensating procedures in place for dealing with different types of data that will be uncovered.

Within the DLP space, there are four major types of technologies.

- **Anti-data leakage**—mainly looks at and filters email. This is the most common solution utilized in an organization to stop ill-advised activities.
- **Port and device control**—good at hardening agents and can look at a computer's registry. This technology is a promising method for identifying rogue wireless networks.
- **Encryption (whole, partial disk, or file-based encryption)**—offers the lowest-hanging fruit. Encryption solves most problems related to lost laptops and other mobile devices. The important consideration factor is key management and device recovery. Many agencies continue to struggle with the Office of Management and Budget's (OMB) data encryption mandate.
- **Database security**—These tools, which are best at protecting structured data, are used in conjunction with the other technologies mentioned above. It's an IANS best practice to layer this technology after a database's inherent security features have been fully utilized.

Today, the DLP vendor/solution is immature and no one product incorporates all of these technologies. However, while existing DLP technologies can't stop targeted malicious attacks, they can be effective at stopping ill-advised employee behavior. ■

About the Author

Allan Carey | is the Senior Vice President of Research and Product Development at the Institute for Applied Network Security (IANS). In this position, he manages all research and intellectual property across the Institute. Prior to IANS, Mr. Carey spent seven years at IDC, a global provider of market intelligence and advisory services for the IT sector. He developed and managed the Security Services practice and provided in-depth analysis, intelligence and consulting on key aspects of the information security and business continuity services markets. He may be reached at the Institute for Applied Network Security, 15 Court Square, Suite 1100, Boston, MA 02108, by telephone at 617/399-8100, or by email at *acarey@ianetsec.com*.

▷ continued from page 23

workshop, entitled "Take Intelligent Risk and Optimized Decisions Based on Time, Available Resources and Risk Tolerance Limits," outlined techniques for adaptive resource management that can be applied to cyber-physical systems.

A manageable CPS handles faults that it encounters regardless if those faults have natural or man-made causes. Fielding a question from the audience that sought a precise boundary between the "C" and the "P" of CPS, Dr. Kwiat responded by citing this paraphrase of an Albert Einstein quote—"Nature is crafty, but not malicious." Dr. Kwiat remarked that since Einstein was a renowned physicist, presumably nature in this regard pertains exclusively to physical and not human nature. Therein is a distinction between physical and cyber—humans influence both but do so more readily in the cyber domain, and human nature, it was noted, can be both crafty and malicious. ■

About the Author

Dr. Kevin A. Kwiat | is a Principal Computer Engineer with the Air Force Research Laboratory, Information Directorate, Rome Research Site where he has worked for over 25 years. He holds 3 patents. He has a BS in computer science and a BA in mathematics from Utica College of Syracuse University and MS and PhD degrees in computer engineering from Syracuse University. His main research interest is the adaptation of concepts from the domain of fault-tolerant computing for information assurance.

FREE Products Order Form

Instructions: All IATAC LIMITED DISTRIBUTION reports are distributed through DTIC. If you are not a registered DTIC user, you must do so prior to ordering any IATAC products (unless you are DoD or Government personnel). To register online: *http://www.dtic.mil/dtic/registration*. The *IAnewsletter* is UNLIMITED DISTRIBUTION and may be requested directly from IATAC.

Name _____ DTIC User Code _____

Organization _____ Ofc. Symbol _____

Address _____ Phone _____

_____ Email _____

_____ Fax _____

Please check one: ☐ USA ☐ USMC ☐ USN ☐ USAF ☐ DoD
☐ Industry ☐ Academia ☐ Government ☐ Other

Please list the Government Program(s)/Project(s) that the product(s) will be used to support: _____

_____

LIMITED DISTRIBUTION

IA Tools Reports (softcopy only)
☐ Firewalls ☐ Intrusion Detection ☐ Vulnerability Analysis

Critical Review and Technology Assessment (CR/TA) Reports
☐ Biometrics (soft copy only) ☐ Configuration Management ☐ Defense in Depth (soft copy only)
☐ Data Mining (soft copy only) ☐ IA Metrics (soft copy only) ☐ Network Centric Warfare (soft copy only)
☐ Wireless Wide Area Network (WWAN) Security ☐ Exploring Biotechnology (soft copy only)
☐ Computer Forensics* (soft copy only. DTIC user code MUST be supplied before these reports will be shipped)

State-of-the-Art Reports (SOARs)
☐ Data Embedding for IA (soft copy only) ☐ IO/IA Visualization Technologies (soft copy only)
☐ Modeling & Simulation for IA (soft copy only) ☐ Malicious Code (soft copy only)
☐ Software Security Assurance ☐ A Comprehensive Review of Common Needs and Capability Gaps
☐ The Insider Threat to Information Systems

UNLIMITED DISTRIBUTION

IAnewsletters Hardcopies are available to order. The list below represents current stock.
Softcopy back issues are available for download at *http://iac.dtic.mil/iatac/IA_newsletter.html*

Volumes 4		☐ No. 2	☐ No. 3	☐ No. 4
Volumes 5	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 6	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 7	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 8	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 9	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 10	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4
Volumes 11	☐ No. 1	☐ No. 2	☐ No. 3	☐ No. 4

Fax completed form to IATAC at 703/984-0773

Calendar

February

IA Symposium
3–6 February 2009
Dallas, TX
http://www.nsa.gov/ia/events/index.cfm

Electronic Warfare Fundamentals and Planning Course
9–13 February 2009
Alexandria, VA

Basic Computer Network Operations Planners Course
23 February–6 March 2009
Alexandria, VA

Phoenix Challenge Conference
24–26 February 2009
Laurel, MD
https://www.phoenixchallengeconf.org

March

Information Processing Interagency Conference 2009
1–5 March 2009
Orlando, FL
https://www.technologyforums.com/9IP

IANS Mid-Atlantic Information Security Forum
3–4 March 2009
Washington, DC
http://www.ianetsec.com

Secure IT 2009 Conference
4–6 March 2009
Los Angeles, CA
http://www.secureitconf.com

Theory of Cryptography Conference (TCC) 2009
15–17 March 2009
San Francisco, CA
http://crypto.stanford.edu/tcc09

AFCEA Belvoir Industry Days
26–27 March 2009
National Harbor, MD
http://www.fbcinc.com/event.aspx?eventid=Q6UJ9A00HG48

April

Wireless Communications and Networking Conference (WCNC)
5–8 April 2009
Budapest, Hungary
http://www.ieee-wcnc.org

DTIC 2009 Conference
6–8 April 2009
Alexandria, VA
http://www.dtic.mil/dtic/announcements/conference.html

Basic Computer Network Operations Planners Course 003
6–17 April 2009
Alexandria, VA

2009 DISA Customer Partnership Conference
20–24 April 2009
Anaheim, CA
http://www.disa.mil/conferences

Information Operations Capabilities Application and Planning Course (IOCAP) 003
27 April–8 May 2009
Alexandria, VA

To change, add, or delete your mailing or email address (soft copy receipt), please contact us at the address below or call us at: 703/984-0775, fax us at: 703/984-0773, or send us a message at: iatac@dtic.mil

IATAC
Information Assurance Technology Analysis Center
13200 Woodland Park Road, Suite 6031
Herndon, VA 20171