U.S. DEPARTMENT OF ENERGY

PNNL-20821

Cyber Friendly Fire

FL Greitzer
TE Carroll
AD Roberts

Prepared for the Air Force Research Laboratory Human Effectiveness Directorate

September 30, 2011

Pacific Northwest
NATIONAL LABORATORY

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
*under Contract DE-AC0*5*-76RL01830*

Printed in the United States of America
Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

PNNL-20821

Cyber Friendly Fire

FL Greitzer
TE Carroll
AD Roberts

September 2011

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Acknowledgments

The authors wish to express appreciation to Dr. Dee Andrews and Patricia Fitzgerald, AFRL Human Effectiveness Directorate, for their support and direction. We also thank JD Fluckiger, PNNL cyber security expert, for providing valuable advice and critique of concepts discussed in this report. The views expressed in this report are the opinions of the Authors, and do not represent official positions of the Pacific Northwest National Laboratory, the Department of Energy, or the Air Force Research Laboratory.

Contents

Summary

Cyber friendly fire (FF) is a new concept that has been brought to the attention of Department of Defense (DoD) stakeholders through two workshops that were planned and conducted by the Air Force Research Laboratory (AFRL) and research conducted for AFRL by the Pacific Northwest National Laboratory. With this previous work in mind, we offer a definition of cyber FF as *intentional offensive or defensive cyber/electronic actions intended to protect cyber systems against enemy forces or to attack enemy cyber systems, which unintentionally harms the mission effectiveness of friendly or neutral forces*. Just as with combat friendly fire, a fundamental need in avoiding cyber FF is to maintain situation awareness (SA). Cyber SA concerns knowledge of a system's topology (connectedness and relationships of the nodes in a system), and *critical knowledge elements* such as the characteristics and vulnerabilities of the components that comprise the system (and that populate the nodes), the nature of the activities or work performed, and the available defensive (and offensive) countermeasures that may be applied to thwart network attacks. To address training requirements and decision support approaches to reduce cyber FF, we examined factors that impact cognition and human performance, particularly with regard to SA. Factors that impact performance include stress, overlearning, and issues relating to cognitive bias. A training implication is to raise awareness and understanding of critical knowledge elements, and a decision support approach is to provide decision aids and/or visualizations that support critical knowledge elements. Mitigation strategies to combat cyber FF – including both training concepts and suggestions for decision aids and visualization approaches – include the following:

- Training recommendations
    - Incorporate stress situations and stress management techniques
    - Develop realistic scenarios that systematically vary stress
    - Address challenges in preparing cyber warriors to overcome cognitive biases
    - Conduct experiments to assess effectiveness of different training approaches

- Information analysis and decision support recommendations
    - Conduct experiments to help identify effective features of decision support and information visualization tools
    - Perform cognitive engineering research to develop prospective information analysis and visual analytics solutions to enhance SA and decrease cyber FF.

To support cyber FF research, we developed an unclassified security test range comprising a combination of virtual and physical devices that present a closed network for testing, simulation, and evaluation. This network offers services found on a production network without the associated costs of a real production network. Containing enough detail to appear realistic, this virtual and physical environment can be customized to represent different configurations.

We conducted an initial pilot study to assess the feasibility of using the test range for controlled experimental studies. Finally, we discuss possible future research directions.

[This Page Intentionally Blank]

1. Introduction

Cyber friendly fire (FF) is a new concept that has been brought to the attention of Department of Defense (DoD) stakeholders through two workshops (in February 2009 and May 2010) that were planned and conducted by the Air Force Research Laboratory (AFRL). The user community and stakeholders are just beginning to grasp and define the concept and its ramifications, and relatively little has been published. To date there have been two published definitions of cyber FF. The first, from Greitzer et al. (Greitzer, Clements et al. 2009), is:

> Cyber Fratricide, or cyber friendly fire, refers to intentional, offensive, or defensive cyber/electronic actions intended to protect cyber systems against enemy forces or to attack enemy cyber systems, which result in inhibiting, damaging, or destroying friendly or neutral infrastructure or operations.

Andrews and Jabbour (Andrews and Jabbour 2011) provide the second:

> The employment of friendly cyber defenses and weapons with the intent of either defending the blue cyber systems from attack from red or gray forces, or attacking the enemy to destroy or damage their people, equipment, or facilities, which results in unforeseen and unintentional damage to friendly cyber systems.

These definitions have many similarities: cyber FF is a consequence of offensive or defensive actions, the actions were performed with purpose, and the damage occurs to friendly or neutral cyber assets. Both definitions imply or overtly identify consequences of the action as unintentional. Furthermore, incidents that are born from accidents, negligence carelessness, or malicious insiders are not friendly fire. From there, the definitions diverge somewhat. Greitzer et al. consider harm to both cyber systems and mission effectiveness, while Andrews and Jabbour focus only on systems. A recent Air Force chief scientist's report on technology horizons mentions the need for "a fundamental shift in emphases from 'cyber protection' to 'maintaining mission effectiveness' in the presence of cyber threats" (United States Air Force Chief Scientist (AF/ST) 2010). Thus, mission effectiveness, and not only systems, is an appropriate focus for friendly fire incidents. In addition, we argue that cyber FF consequences may be felt well beyond cyber space. Consider cyber physical systems that closely integrate physical, computational, and communication components to sense and effect changes in the real world. These systems are heavily employed in critical infrastructure to control and monitor processes. Adversely impacting the operation of these systems may result in large-scale power failures, toxic waste releases, or explosions that can have catastrophic consequences on the environment and life.

With this discussion in mind, we offer the following revised definition of cyber FF:

> *Cyber friendly fire is intentional offensive or defensive cyber/electronic actions intended to protect cyber systems against enemy forces or to attack enemy cyber*

> *systems, which unintentionally harms the mission effectiveness of friendly or neutral forces.*

The following two examples illustrate cyber FF incidents that derive from defensive actions that unintentionally harm the organization's missions:

- **Illustrative Example #1.** Company XYZ moved their corporate website and email to a hosting company to save money. A hacker who obtained an administrator's account on the hosting company's servers begins to disrupt services by attempting to hack into Company XYZ's hosted web server. An administrator at Company XYZ notices this hacking activity and quickly takes actions to protect company resources by blocking traffic from the hosting company. Company XYZ is no longer able to obtain access to their corporate website or their email, which resides at the hosting company.

- **Illustrative Example #2.** A current vulnerability to widely-deployed web serving software is being actively exploited. The vendor for the software has issued a security patch. Company ABC, who relies on the software as a critical component of their e-business platform, rapidly deploys the fix on their infrastructure. The patched software exhibits abnormal behavior and causes transactions to fail.

The next examples illustrate defensive actions that harm friendly assets, but do not constitute FF:

- **Illustrative Example #3.** Company XYZ stores client personally identifiable information in a central database. The database is compromised by an adversary, who then actively engages in exfiltrating the stored data. Company XYZ administrators detect the extrusion of data and take action to stem the flow of data by severing the Internet connection until they can remediate and recover from the attack. The administrators fully comprehend that no client is able to access the company's services while disconnected, but the induced harm is far less than harm of continued data exfiltration.

- **Illustrative Example #4.** A network administrator is writing a new firewall rule to block specific malicious network traffic. Before the rule can be completed, the administrator's Bengal cat leaps onto her keyboard, depressing several keys, which mangles and activates the rule. The rule disrupts traffic to the company's web server cluster, inhibiting clients from processing products.

2. Cognitive Approaches to Cyber Friendly Fire Research

The concept of cyber FF is similar in many respects to combat friendly fire (Greitzer, Clements et al. 2009), and from a cognitive perspective, the fundamental issue is with maintaining situation awareness (SA). The scientific literature on SA is substantial and no attempt is made here to report exhaustively on this topic. In short, the most accepted definition of SA is given by Endsley (Endsley 1995): SA is the perception of the elements in the environment within a volume of time and space (Level 1 SA), the comprehension of their meaning (Level 2 SA), and the projection of their status into the future (Level 3 SA).

SA depends on an accurate mental model (Sarter and Woods 1991). Mental models have been described as well-defined, highly organized, and dynamic knowledge structures that are developed over time from experience (e.g., (Kozlowski 1998)). By representing organized "chunks" of information in the environment, mental models serve to reduce the information load that would otherwise overwhelm the ability of decision makers to attend, process, and integrate the large amount of information that is inherent in complex operational environments. Cues in the environment activate these mental models, which in turn guide the decision-making process. Appropriate and effective mental models enable experienced decision makers to correctly assess and interpret the current situation (Level 1 and Level 2 SA) as well as to select an appropriate action based on patterns (mental models) stored in their long-term memory (Serfaty, MacMillan et al. 1997).

2.1 Cyber Situation Awareness

Considering that a lack of SA is often a contributing factor to human errors in decision making, it is clear that a study of cyber FF should focus on factors that affect the cyber security officer's/system administrator's SA. What constitutes *cyber* SA?

Tada and Salerno (2010) mapped constructs of SA to more cyber-relevant network environments. A SA process model was constructed that has general applicability as well as specific relevance to cyber SA. The paper also suggested a set of metrics that may be useful in assessing the effectiveness of tools for supporting SA. Consistent with Tada and Salerno's characterization of SA, our notion of cyber SA focuses on knowledge of a system's topology (connectedness and relationships of the nodes in a system), the characteristics and vulnerabilities of the components that comprise the system (and populate the nodes), the nature of the activities or work performed, and the available defensive (and offensive) countermeasures that may be applied to thwart network attacks. SA must also include an understanding of *why* each node exists, *what* it is doing, and the harm associated with disrupting that function as a response to attack. The trade-offs between accepting the ongoing risks of attack must be properly balanced against the damage done to the overall organization's mission, and the process of balancing those elements should motivate and guide the defender to select responses that minimize the total amount of harm.

More specifically, we may speculate on implications for cyber defense and cyber SA based on the notion of "digital SA."[1] Given the complexity of cyber structures (particularly at the national scale of critical infrastructures such as the Internet or the electric power grid), it is necessary to take a "system of systems" perspective. In this view, there is never 100% certainty or complete knowledge, and it must be assumed that systems will be attacked (i.e., it is not possible to prevent all attacks with certainty). Thus, an appropriate cyber security strategy is *resiliency*, i.e., the ability to anticipate, avoid, withstand, minimize, and recover from the effects of attacks (or for that matter, from the effects of natural disasters). To anticipate and avoid the effects of attacks or other adverse circumstances, a high level of SA is required. In particular, there is a critical need for operators to *anticipate* and *apply protocols* to avoid *cascade effects* in the network, thereby avoiding unintended consequences of defensive or offensive actions. The following types of knowledge (*critical knowledge elements*) are required to invoke this anticipatory process:

- Knowledge of each enterprise/component of the network
- Knowledge of each computer system of interest in each enterprise/component
- Knowledge of each I/O port on each computer and how it is being used
- Record of traffic flow and volume on every I/O port
- Knowledge of the results of computing expected during the normal operation of each of the components in the network based on the current traffic flow and volume
- Knowledge of operating limits for each component, enabling the decision maker to project "faults" that may lead to shut-downs and cascade failures
- Knowledge of alternative corrective actions for such faults.

An additional consideration regarding the role of SA and cognitive models in cyber FF is the importance of Team SA: the degree to which each team member possesses the SA required for his or her responsibilities (Endsley 1995) and in particular, the extent to which team members possess the same SA on *shared* SA requirements (Endsley and Jones 1997; Endsley and Jones 2001). Conflicts between goals and/or failures to coordinate goals among different members of the team are major underlying/root causes of many cyber FF incidents.[2]

Given these considerations, a recommended approach to study SA and cyber FF is to adopt a cognitive systems perspective, and particularly a naturalistic decision making approach, to capture the mental models that constitute the above types of knowledge. Implications for training and/or tool development include:

- A strategy to train operators within this naturalistic decision-making paradigm to raise awareness and understanding of the above critical knowledge elements.

[1] The following discussion is based in part on an essay on situation awareness in Wikipedia: http://en.wikipedia.org/wiki/Situation_awareness

[2] The scenario in Section 3.2, above, provides a typical example: The network administrator's goal of protecting the network from general degradation resulting from the denial of service attack is at odds with the Intensive Care Unit's goal of closely monitoring patients' status.

- A tool development strategy to design and implement decision aids and/or visualizations that support the acquisition of, or use of, the above critical knowledge elements.

2.2 Four Trends That Make Digital SA Harder

Four current trends greatly increase the difficulty of performing digital SA.
(1) First, missions are defined in terms of abstract resources and not actual systems and devices. For example, a mission in support of business-to-business portal is defined in terms of number of concurrent users and user experience attributes, such as page response time. The requirements are translated into resource and location requirements (e.g., "ten web servers in the East Coast data center will be tasked for this mission"). The mission planner may never be aware of what actual resources are allocated, the underlying network topology, or where the resources are even geographically located.

(2) The second trend that makes digital SA hard is that organizations are outsourcing the responsibility of infrastructure to third party providers who build and maintain an independent infrastructure that concurrently supports one or more autonomous organizations. The provider may not be external—it may be a separate department within the organization that supports all the organization's missions (e.g., Defense Information Systems Agency, DISA, supports the IT infrastructure for the DoD). Infrastructure As A Service (IAAS) exemplifies this practice. The provider may, at its discretion and in compliance with Service Level Agreements (SLAs), perform maintenance that may temporarily disrupt service; depending on the sensitivities of missions, this may reduce effectiveness. Combined with Trend 1, communication between the parties is difficult because organizations speak/plan in terms of missions while the provider speaks/plans in terms of resources.

(3) Dynamic management of resources, as seen in cloud, grid, and utility computing environments, make for flexible resource allocations that are revised with changing demands and requirements. Even if the users are certain about the identity of the resources at a time *t*, the cloud management may choose to migrate at time *t*+1 the processes to systems made up of different components that exist in a different continent. And of course, the underlying network topology is different too.

(4) Finally, organizations are augmenting their networks with increasingly large number of sensors, which, as one may expect, is overloading human analysts with oceanic volumes of data. The theory behind this trend is that by capturing all information available at the device and network level that it should make possible of detecting of any and all attacks. Unfortunately, there has not been a corresponding improvement in data fusion, analysis, and detection methods, and the vast amounts of data have swamped analysts.

[This Page Intentionally Blank]

3. Test Bed to Support Experimentation

Experimentation involves designing several test scenarios in embodying the cyber security operations of a test domain that should evoke cyber FF incidents. In our specific case, the test domain is a managed server provider who offers web-enabled applications to the general public. This domain was selected because of its familiarity to our available participant pool, which includes college student interns, cyber security operators, and analysts at Pacific Northwest National Laboratory. We investigate the test scenarios using a dual-task experimental framework that places participants in situations that better reflect real-world conditions. In this paradigm, a participant performs a *primary task* while being occupied by one or more *secondary tasks*. As an example, a participant is instructed to review system logs (a secondary task) while defending against a network-borne cyber-attack (the primary task). The participant's performance in these tasks reflects the amount of "cognitive load" imposed by one or the other. Increasing the difficulty (e.g., complexity, tempo, etc.) of the primary task will decrease the participant's cognitive capacity to perform secondary tasks. Thus, arranging cyber security tasks in controlled environments supports the study of cognitive factors underlying SA and resultant FF incidents. Applicable metrics to investigate cognitive load include completion time of the primary and/or second tasks, how successful the participant was in performing her tasks, and the quality of the work.

A recommended venue in which to conduct the experiment would be PNNL's unclassified security test range, a test bed that can be arbitrarily configured so to simulate a multitude of different organization's networks.

3.1 The Unclassified Security Test Range

The unclassified security test range consists of a combination of virtual and physical devices to create a closed network for testing, simulation, and evaluation. This closed network offers services found on a production network without the costs associated with duplicating a production network. The idea is to duplicate enough of a real network to allow the closed network to appear realistic.

The virtual and physical environment can be customized to represent different configurations based on requirements. For example, it is possible to create simulated environments that emulate Fortune 500 enterprises, and application and infrastructure service providers.

3.1.1 Test Range Specification

The test range utilizes VMware ESX technology to provide Virtual Machines for the following:

- Hundreds of Windows XP Professional workstations
- Microsoft Windows 2008 Servers
- Ubuntu Linux Desktop or Servers
- Microsoft Active Directory / Domain Controller Servers
- Microsoft Exchange Mail Servers

- Domain Name Services
- SQL Database Servers
- SysLog Servers
- Microsoft IIS HTTP Servers
- Apache HTTP Servers

Each of the services listed above has a baseline configuration created and is backed up to an external backup source in the form of an image. This image allows the restoration of a known configuration to its original state. This may occur for several reasons, including: at the end of a simulation, test, or if any of the services in the range fail or are compromised. The restore time for any device is an average of 15 minutes with the use of Clonezilla imaging software. It is also possible to create a snapshot of virtual devices in case a need arises to test a particular feature without the need for a full system restore. This also allows for the analysis of a specific state or point in the simulation. If a service not listed above is needed, it can be added to the range by request.

The test range has a monitoring feature that provides the capability to monitor, log, and analyze all the traffic flowing through the network. This capability is made available using a Cisco layer 3 switch with the SPAN port feature, which allows the switch to duplicate all traffic being sent across all switch ports to a single port connected to a monitoring workstation. The monitoring workstation can run any type of software capturing program depending on the requirements. The typical software tools used for this capability in the test bed is Wireshark and TCPdump.

Another software package, called ANTS, is available on the test range for creating real user traffic on the virtual test network to enable simulations in what appears to be a normal live network. This software package allows for the creation of white noise on the network to aid in simulating a production network. This software package also allows for the creation of white noise on the network to aid in simulating a production network. It is scalable and can simulate up to 200 unique devices on each server in the virtual network, each with its own unique IP and MAC address. The devices can mimic employees performing their job on their computers, or it can launch a malicious attack against a target device. This capability also saves on costs since it allows the range to simulate traffic without the need for extra physical computer workstations or large servers.

The SEAL software package installed on the test range provides a remote view into a closed range. This is done by installing a dual network interface card in each range resource, which allows communication with a SEAL server in addition to providing normal network communications. The SEAL server acts as a boundary between the closed network range and the outside network. The only communication allowed between the SEAL server and the outside network is done via a secure TCP port. This secure TCP port allows a SEAL client to remotely view each virtual server and its connected desktops or workstations and interact with them. Without this capability, the operator would need to interact with each piece of equipment from a local console instead of interacting with it from his/her office or another location.

3.1.2 Hardware Specification

The hardware used in the range consists of Servers, physical workstations, switch equipment, firewalls, and remote management devices. Servers and Network Equipment that comprise the **Cyber Friendly Fire Test Bed** include:

- Ten Dell PowerEdge 2950 servers with dual 3.2 GHz processors and 16 gigabytes of memory
- Dell PowerEdge R900 servers with dual 3.2 GHz processors and 8 gigabytes of memory
- Fifty iGoLogic embedded computers with rack mount capability
- Cisco 3750E network switches
- Cisco 3560G network switches
- Cisco Adaptive Security Appliances capable of acting as a firewall or VPN appliance

Figure 1 shows an example of the network diagram configuration.

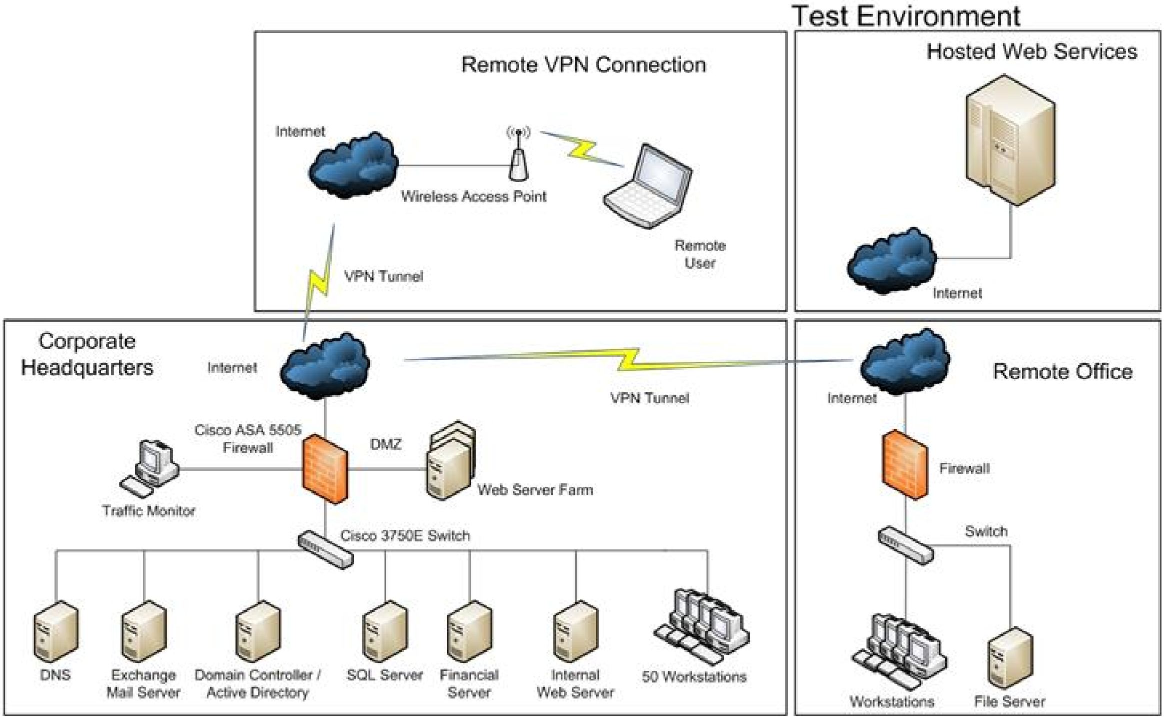


Figure 1. Logical Network Diagram of the Test Environment Baseline

3.1.3 Instrumentation for Performance Data Collection

Performance data are collected through the monitoring feature described in Section 2.1 and can be stored in any desired file format. The default storage method involves creating a **libpcap** packet capture file. The type of information collected with this method includes the time stamp, the protocol type, the internet protocol header, the packet data, source IP address, and the destination IP address. Using protocol analysis tools such as **Wireshark** or **ntop**, it is possible to determine the start and end time of a simulation and identify what events are occurring on the network. Visualization tools can also be used to show a real time SA view of the network.

3.2 Experimental Setup

The test range was configured to appear as an Internet-connected Managed Service Provider (MSP) offering specialized web applications to the general public at large. The test range has many components that one would observe in a small- to medium-sized organization, including Microsoft Active Directory authentication and authorization servers, Microsoft Exchange mail servers, Microsoft SQL Server database servers, IIS web servers, and network sensors. Figure 2 is a logical network diagram of the setup. The network is essentially divided into a production component that hosts the web and network services, and a user component that hosts thirty employee workstations and other end devices. The organization's network is separate from the Internet by a Cisco ASA network security device that both firewalls and detects intrusions.

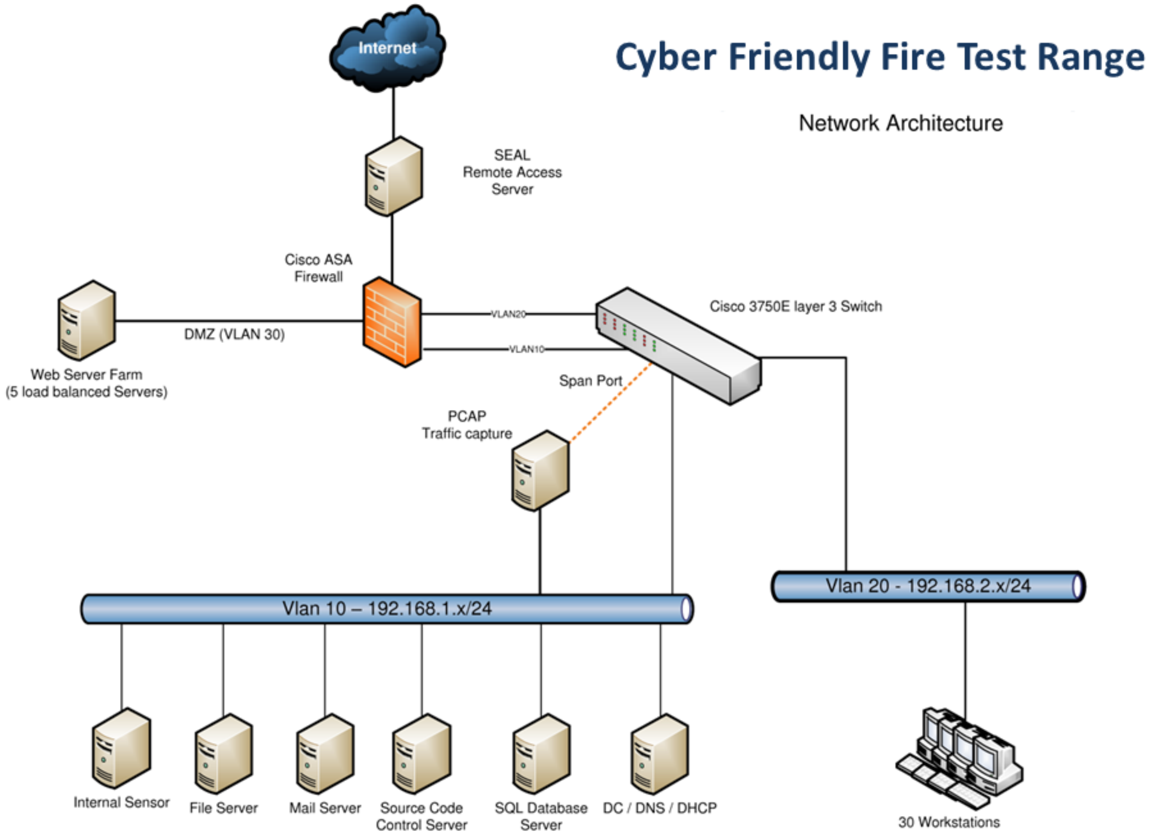


Figure 2. The network diagram for the experiments

Business sensitive information is stored on the File Server, Mail Server, Source Code Control Server, and SQL Database Server. The File Server houses thousands of internal documents, such as finance and technical designs. The Mail Server stores messages between all of the organization’s employees, including the CEO, CFO, and CIO. All of the organization’s source code is saved in the Source Code Control Server. Personally Identifiable client data are saved in the SQL Database Server. Releasing of any of this information to unauthorized parties would have a significant, detrimental impact on the organization’s reputation, which would harm earnings. In many states, the unauthorized release of Personally Identifiable Information (PII) requires reporting the incident to the government and the purchasing of monitoring and insurance plans to indemnify the financial losses to the aggrieved persons and businesses. The valuable information stored in these servers pose obvious points of interest for an adversary.

To support studies in cyber SA and cyber FF, the following scenarios have been developed and programmed into the test range:

- **Scenario #1.** The network operator notes a high page request volume on company XYZ’s WWW server cluster. Customers have phoned the company, reporting problems about slow application response times and “hung” connections. The operator needs to determine if this is a Distributed Denial of Service (DDoS) attack launched by one of many freely available tools or is the result of potential client interest in an improved application that was recently released.
- **Scenario #2.** US Cert announces a zero-day vulnerability for company XYZ’s web hosting platform. Hundreds of messages in cyber security-related forums suggest that attackers were surreptitiously exploiting the fault for several months prior. XYZ fits the current profile of companies that have been targeted. The platform’s vendor has released a patch. While the patch does fix the vulnerability, it may introduce other faults that cause other services to fail.
- **Scenario #3.** Indicators are activated for unusual traffic and accesses to the SQL Server Database from developer workstations. The database plays an integral role in the application operations. It is quite obvious to clients that the database is down. Business reputation would be greatly harmed if data is released to unauthorized parties.

[This Page Intentionally Blank]

4. Mitigation Approaches

4.1 Training

To address training requirements and approaches to reduce cyber FF, it is useful to examine factors that impact cognition and human performance, particularly with regard to SA. Research has demonstrated a number of factors that impact performance; in the present context, effects of stress, overlearning, and issues relating to cognitive bias are particularly relevant. Greitzer and Andrews (2010) review cognitive foundations and implications for training to mitigate combat friendly fire. Here we describe aspects of this research that are pertinent to training requirements for cyber FF.

4.1.1 Effects of Stress on Performance

Stress has strong effects on every aspect of cognition from attention to memory to judgment and decision making. Under stress, attention appears to channel or tunnel, reducing focus on peripheral information and centralizing focus on main tasks (Kavanagh, 2005). Originally observed by Kohn (1954), this finding has been replicated often, first by seminal work from Easterbrook (1959) demonstrating a restriction in the range of cues attended to under stress conditions (tunneling) and many other studies (see Staal, 2004). Research by Janis and Mann (1977) suggests that peripheral stimuli are likely to be the first to be screened out or ignored, and that under stress, individuals may make decisions based on incomplete information. Similarly, Friedman and Mann (1993) note that individuals under stress may fail to consider the full range of alternatives available, ignore long-term consequences, and make decisions based on oversimplifying assumptions—often referred to as heuristics.[3] Research on the effects of stress on vigilance and sustained attention, particularly regarding effects of fatigue and sleep deprivation, shows that vigilance tends to be enhanced by moderate levels of arousal (stress), but sustained attention appears to decrease with fatigue and loss of sleep (Davies and Tune, 1970).

4.1.2 Overlearning

Several investigations have shown that tasks that are well-learned tend to be more resistant to the effects of stress than those that are less-well-learned. Extended practice leads to commitment of the knowledge to long term memory and easier retrieval, as well as automaticity and the proceduralization of tasks. These over-learned behaviors tend to require less attentional control and fewer mental resources (Leavitt, 1979; Smith & Chamberlin, 1992), which facilitates enhanced performance and yields greater resistance to the negative effects of stress—i.e., overlearned behaviors are less likely to be forgotten and more easily recalled under stress. Van Overschelde and Healy (2001) found that linking new facts learned under stress with preexisting knowledge sets helps to diminish the negative effect of stress. On the other hand, there is also a tendency for people under stress to “fall-back” to early-learned behavior—even less efficient or more error prone

[3] While researchers who argue that perceptual narrowing reduces the quality of individual decisions, Klein (1996) observed that the use of heuristics may allow individuals to respond more quickly to external demands while under stress or when provided only partial information.

behavior than more recently-learned strategies—possibly because the previously learned strategies or knowledge are more well-learned and more available than recently acquired knowledge (see Allnut, 1982; Barthol & Ku, 1959; Zajonc, 1965).

4.1.3 Effects of Stress on Learning

Research suggests that high stress during instruction tends to degrade an individual's ability to learn—perhaps due to interference or disruption in the encoding and/or maintenance phases of working memory. An implication for instructional strategies is that a phased approach should be used, with an initial learning phase under minimum stress, followed by gradual increasing exposure to stress more consistent with real-world conditions. Stress inoculation training attempts to immunize an individual from reacting negatively to stress exposure. The method provides increasingly realistic pre-exposure to stress through training simulation; through successive approximations, the learner builds a sense of positive expectancy and outcome and a greater sense of mastery and confidence. This approach also helps to habituate the individual to anxiety-producing stimuli.

4.1.4 Team Performance

Finally, it is important to consider group processes in this context. Research on team decision making indicates that effective teams are able to adapt and shift strategies under stress; therefore, team training procedures should teach teams to adapt to high stress conditions by improving their coordination strategies. Driskell, Salas, and Johnston (1999) observed the common phenomenon of Easterbrook's attentional narrowing is also applicable to group processes. They demonstrated that stress can reduce group focus necessary to maintain proper coordination and SA—i.e., team members were more likely to shift to individualistic focus than maintaining a team focus.

4.1.5 Implications

Based on the foregoing discussion, we can summarize the challenges and needs for more effective training in general terms as well as more specifically focused on cyber defense and mitigation of cyber FF: training should incorporate stress situations and stress management techniques, development of realistic scenarios that systematically vary stress (e.g., as produced by varying cognitive workload through tempo of operations and density of attacks), and addressing challenges in preparing cyber warriors to overcome cognitive biases. The following factors should be included in designing training approaches:

- Training should provide extended practice, promoting more persistent memory and easier retrieval, and to encourage automaticity and the proceduralization of tasks to make them more resistant to the effects of stress.
- Training scenarios should include complex/dynamic threats that reflect the uncertainties of the real world—scenarios that force trainees to operate without perfect information and that incorporate surprises that challenge preconceptions or assumptions.

- Training scenarios should be designed to encourage the habit of testing one's assumptions to produce more adaptive, resilient cyber defense performance in the face of uncertainty.
- Training should enhance awareness of the effects of stress on cognitive performance—such as tunneling and flawed decision making strategies that ignore information—and coping strategies to moderate these effects. The training should be designed to make as explicit as possible what might happen to skill and knowledge under stress.
- Train awareness of cognitive biases and practices for managing these biases
- Emphasize habits of testing assumptions and moving beyond traditional reactive behaviors to train techniques for more adaptive, resilient performance in the face of uncertainty.
- Team training should focus on strategies for maintaining group cohesion and coordination, mitigating the tendency for team members to revert to an individual perspective and lose shared SA.
- Training should exercise the execution of cognitive tasks by both individuals and groups.

To meet these objectives, we have designed the **Cyber Friendly Fire Test Bed** to enable the systematic manipulation of scenario parameters relating to cognitive workload, and we have begun to examine possible performance measures to assess the effects of these independent variables on situation awareness and decision making.

4.2 Tools

A key objective in the study of factors influencing cyber FF and mitigation strategies is to identify features of decision support tools with potential to reduce the occurrence of cyber FF. Our review of relevant research, as summarized in the foregoing discussion, strongly suggests that tools and visualizations to improve cyber SA are key ingredients of desired solutions. Important functions should include decision aids to support memory limitations, to counteract the negative effects of stress on performance (e.g., perceptual narrowing), and to avoid the negative consequences of cognitive biases on decisions.

4.2.1 Supporting Memory Limitations that Reduce Situation Awareness

As stated earlier, support for the cyber analyst should strive to encourage proactive decision making processes that *anticipate* and *apply protocols* to avoid *cascade effects* in the network, and concurrently avoid unintended consequences of defensive or offensive actions. We identified a set of critical knowledge elements required for enhanced SA and anticipatory decision making, including knowledge of components of the network, details of each computer system, I/O ports, traffic flow/volumes, and ability to project impacts of possible courses of action. Decision aids and/or visualization support is needed to alleviate memory lapses and limitations by providing readily accessible information on network topology and component assets/vulnerabilities – typically referred to as external representations or external memory by researchers advocating the study of "distributed cognition" in the broader context of the social and physical environment that must be interwoven with the decision maker's internal representations (also referred to as "situated cognition"; Norman, (1988); Hollan et al., 2000). Thus, a decision aid that displays critical knowledge elements for components that are being considered for application of remedial actions may help to avoid cyber FF effects that impair system effectiveness. This concept is similar to what Tadda and Salerno (2010) refer to as "Knowledge of Us" (data relevant to the importance of assets or capabilities of the enterprise)—hence, a process that identifies to the decision maker whether there is a potential or current impact to capabilities or assets used to perform a mission. Similarly, a tool may be envisioned that helps the decision maker understand and prioritize *risks* that may be computed for various possible alternative actions. An example of a research prototype that might be applied to this problem is the Cyber Attack Risk Inference Model (CARIM).[4]

4.2.2 Mitigating Cognitive Biases

Gestalt psychology tells us that we tend to see what we expect to see. Expectancy effects can lead to such selective perception as well as biased decisions or responses to situations in the form of other cognitive biases like confirmation bias (the tendency to search for or interpret information in a way that confirms one's preconceptions) or irrational escalation (the tendency to make irrational decisions based upon rational decisions in the past). The impact of cognitive biases on decision performance—particularly response selection—is to foster decisions by individuals and teams that are based on prejudices or expectations that they have gained from information learned before they are in the response situation. Decision aids and visualizations are needed that help to reduce confirmation bias, irrational escalation, and other forms of impaired decision making. One possible form of decision support designed to counteract these biases is the use of the analysis of competing hypotheses (e.g., Heuer, 1999). Other concepts that may serve as sources of ideas and strategies for the design of decision aids may be derived from problem solving techniques discussed by Jones in The Thinker's Toolkit (1998).

4.3 Summary of Research Needs

Based on the foregoing discussion, we summarize the challenges and needs for more effective training and decision support to improve cyber defense and mitigate cyber FF:

- Training recommendations
    - Incorporate stress situations and stress management techniques
    - Develop realistic scenarios that systematically vary stress (e.g., as produced by varying cognitive workload through tempo of operations and density of attacks)
    - Address challenges in preparing cyber warriors to overcome cognitive biases
    - Conduct experiments to assess effectiveness of different training approaches

- Information analysis and decision support recommendations
    - Conduct experiments to help identify effective features of decision support and information visualization tools. Will conventional training approaches to improve

[4] See http://i4.pnnl.gov/focusareas/carim.stm.

analytic process (e.g., analysis of alternative hypotheses, other decision making tools and strategies) be effective in the cyber domain? Our intuition suggests that the answer is "no" because of the massive data, extreme time constraints requiring near real-time responses, and the largely data-driven nature of the problem. New types of data preprocessing (triage) and visualization solutions will likely be needed to improve SA.

- Perform cognitive engineering research to develop prospective information analysis and visual analytics solutions to enhance SA and decrease cyber FF.

[This Page Intentionally Blank]

5. Pilot Study

We conducted a small pilot study using PNNL’s Unclassified Test Range test bed as the simulation environment. The purpose of the study was to “test drive” the simulation environment and assess the feasibility of its use for experiments to investigate factors affecting cyber FF. Due to the small sample size of four, no statistical inferences may be made, but we nevertheless treated the study as we would a formal experiment; we also obtained informal feedback from the study participants. Here we provide a brief review of the experiment, method, and results obtained. Approval in advance for this study was obtained from the PNNL Institutional Review Board (PNNL IRB No. 2011-26).

The purpose of the experiment was to examine the effect of providing a visualization aid for cyber data monitoring on possible occurrence of cyber FF.

5.1 Test scenario and Visualization Aid

The simulated environment and basic scenario were as described earlier in this report (Section 3.2). **Error! Reference source not found.** is a diagram of the scenario network. Each participant plays the role of a network/system administer for Paper Clip Inc., a small firm that distributes paperclips and other fasteners worldwide. The operators are responsible for maintaining
the operation of the organization’s small network of one Cisco AS5505 firewall, one Cisco 3710 network switch, one Microsoft Exchange messaging and communications server, one Microsoft SQL Server ODBMS server, one Active Directory server (that also

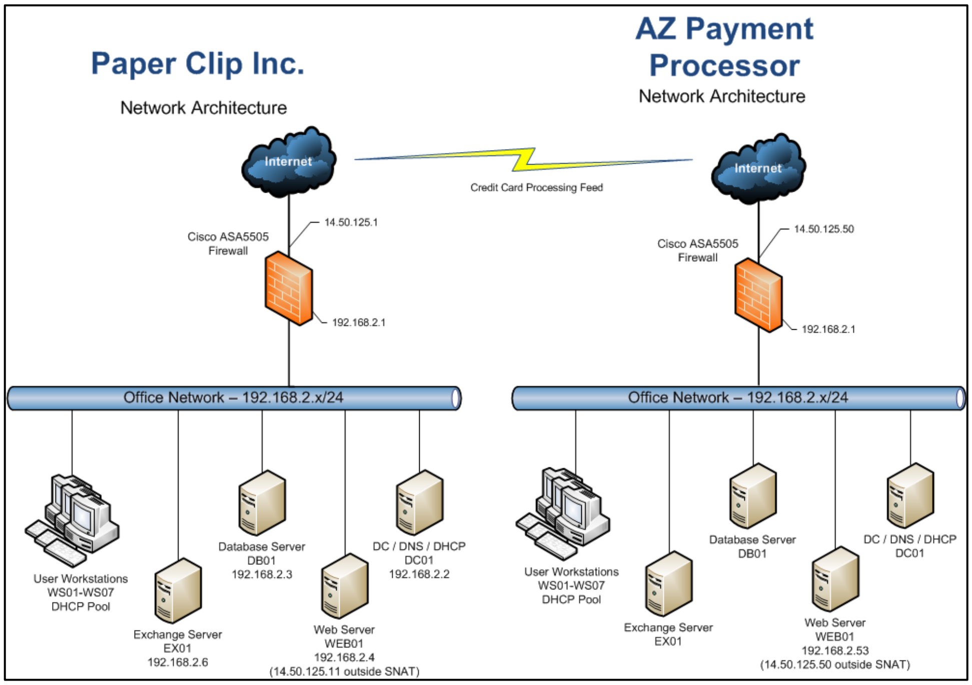
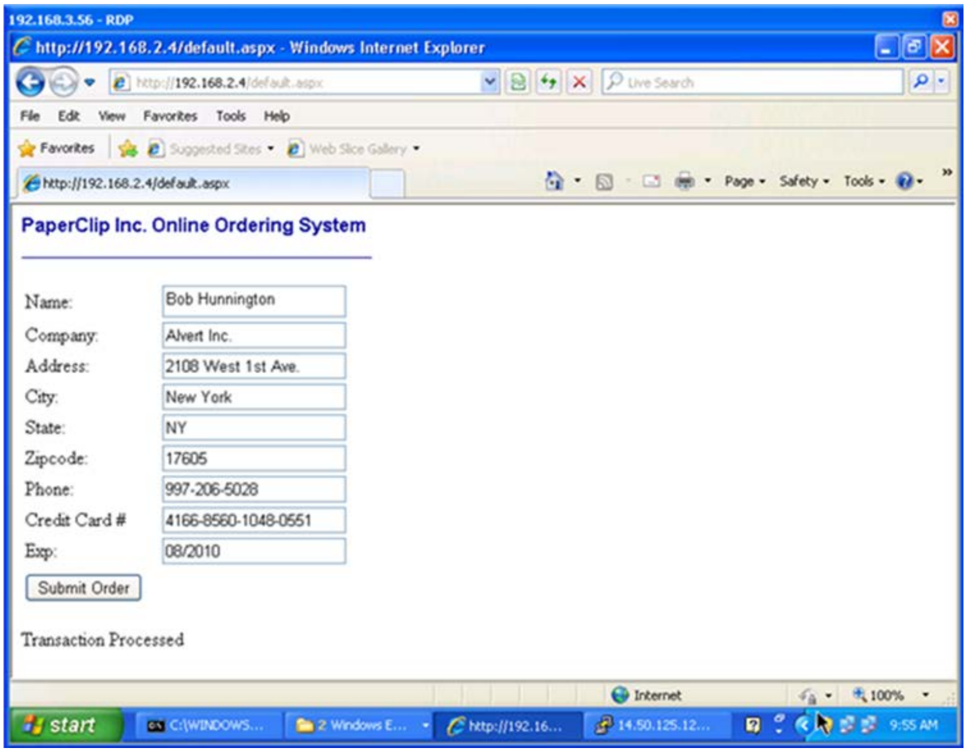


Figure 3. The network diagram of the test scenario. Two organizations are represented, Paper Clip Inc. and AZ Payment Processor. The participant performs in the role of network/system administrator for Paper Clip Inc.

runs an DHCP service to configure the network attributes of the workstations and a DNS service for internal domain name advertising and external domain name navigation), seven Windows XP workstations, and one Microsoft IIS 6 WWW (HTTP) server. A critical component of the business is its web-based order processing application (see Figure 4). Maintaining the operation of this system is critical for the well-being of the company. The application depends on the WWW server, internal ODBMS system, and a credit card payment processing service provided by AZ Payment Processor, a globally renowned payment processor. When a customer places an order, database records are inspected and updated, and the credit card payment information is transferred to AZ Payment Processor to process payment.



PaperClip Inc. Online Ordering System

Name:	Bob Hunnington
Company:	Alvert Inc.
Address:	2108 West 1st Ave.
City:	New York
State:	NY
Zipcode:	17605
Phone:	997-206-5028
Credit Card #	4166-8560-1048-0551
Exp:	08/2010

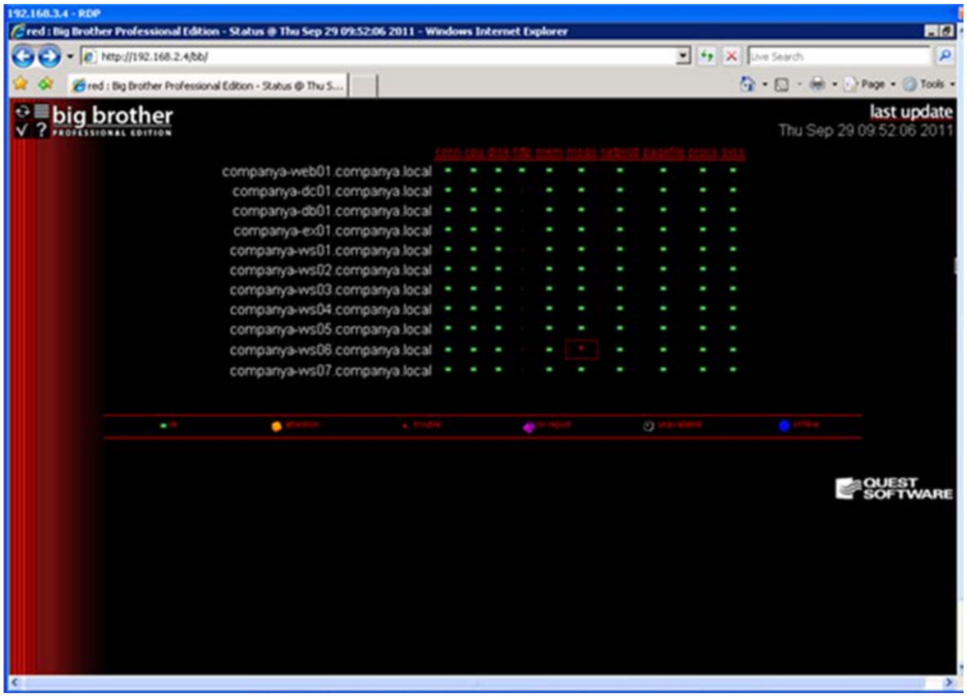
Submit Order

Transaction Processed

Figure 4. A web page from Paper Clip Inc.'s order processing system. The collected credit card information is transferred to AZ Payment Processor to process payments.

During the scenario, we arranged that the participants would have to confront two types of events that harmed the business operations of Paper Clip, Inc. Both events interfered with customers accessing the order processing system. The first event type, which manifested several times during the scenario, was a fault in the order processing system that triggered the abnormal execution termination of the ordering system and the WWW service. The participants were required to reset the WWW service and restart it. The second event was a Denial-of-Service (DoS) that originated from the AZ Payment Processor against the ordering system. While a partner attacking another seems exceptional, there have been cases in which attackers have exploited partner relationships and used compromised partners as stepping stones to further their progress towards their goal. The attack consumed large quantities of resources, slowed the ordering system, and generally resulted in poor performance that can be readily observed by customers. As experimenters, we controlled the start of the events in the scenario. Assuming the role of a help desk, we would inform the participants after one minute about the occurrence of an event.

Every participant was furnished with several commercially available monitoring tools. The first is Big Brother (BB) system and network monitor. The first is Big Brother (BB) system and network monitor. BB is configured to monitor various aspects of system and network object attributes (e.g., CPU utilization, data rate, system event logs) and alerts when these object attributes exceed defined thresholds. BB supplies alert notifications in an easily understood panel. Figure 5 is a screen shot of the simulation's Big Brother network overview that is displaying "all conditions clear." The single alert informs the administrator that the system is unable to download updated malware/virus signatures. By design, the Test Range is isolated and constitute systems are unable to communicate with systems on the Internet. The second tool for monitoring is the Cisco ASA's ASDM panel. The overview panel displays current network conditions, such as data rate and connection volume. Other ASDM panels display detailed network traffic traces and assist in traffic inspection. Half the participants were also furnished with EtherApe,[5] a network monitor that displays network activity graphically. It colorfully renders communication between systems by drawing a link between systems. The width of the link changes in proportion to the volume of traffic, i.e., the link width expands as traffic increases.


192.168.3.4 - RDP
red : Big Brother Professional Edition - Status @ Thu Sep 29 09:52:06 2011 - Windows Internet Explorer
http://192.168.2.4/bb/
Live Search
red : Big Brother Professional Edition - Status @ Thu S...
Page
Tools
big brother
PROFESSIONAL EDITION
QUEST SOFTWARE

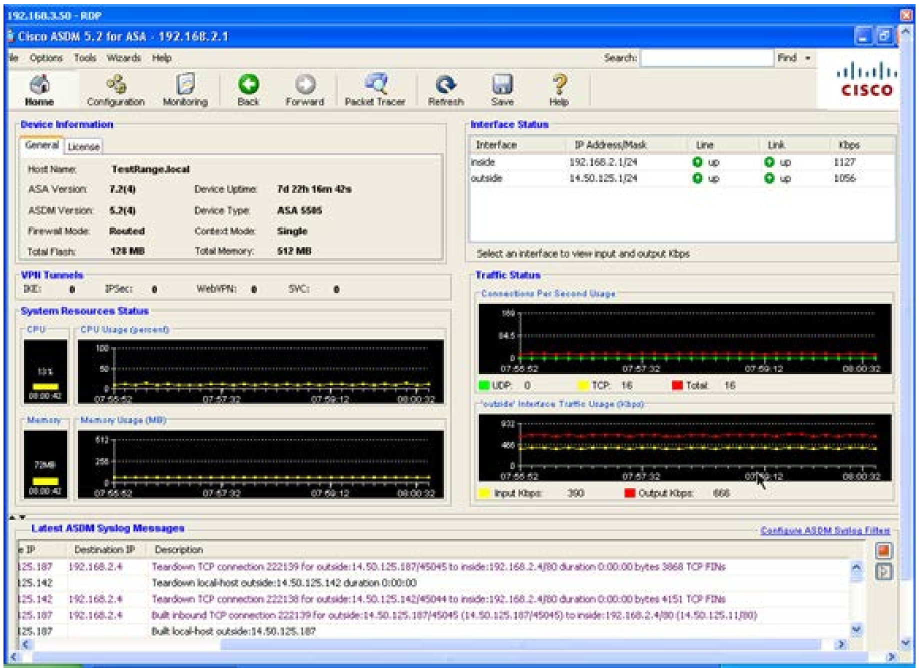
last update
Thu Sep 29 09:52:06 2011

companya-web01.companya.local
companya-dc01.companya.local
companya-db01.companya.local
companya-ex01.companya.local
companya-ws01.companya.local
companya-ws02.companya.local
companya-ws03.companya.local
companya-ws04.companya.local
companya-ws05.companya.local
companya-ws06.companya.local
companya-ws07.companya.local

Figure 5. The Big Brother network and system monitor overview page. This page provides a summary of the health of the monitored network and systems.

[5] EtherApe can be downloaded from http://etherape.sourceforge.net/ (last accessed on September 29, 2011).

We inserted commands into Paper Clip, Inc.'s order system to control its operation. One command resulted in both the ordering system and the WWW service to abnormally terminate. Figure 6 and Figure 7 show ASDM and EtherApe when the network and systems are not experiencing any difficulty, and Figure 8 and Figure 9 show the monitors when the WWW service has abnormally terminated. The general ASDM display shows a decrease in the data rate and increase in the number of connections (when the WWW service fails, the clients repeatedly attempt connection; the connection operations take less time than connection and transfer of data). EtherApe indicates a greater number of links and the width of the links are thin. Unfortunately due to unforeseen interaction, the BB alert displays are dependent on the WWW service, so they fail when the WWW service terminates.



Host Name: TestRange.local

ASA Version: 7.2(4) Device Uptime: 7d 22h 16m 42s

ASDM Version: 5.2(4) Device Type: ASA 5505

Firewall Mode: Routed Context Mode: Single

Total Flash: 128 MB Total Memory: 512 MB

IKE: 0 IPSec: 0 WebVPN: 0 SVC: 0

Interface	IP Address/Mask	Line	Link	Kbps
inside	192.168.2.1/24	● up	● up	1127
outside	14.50.125.1/24	● up	● up	1056

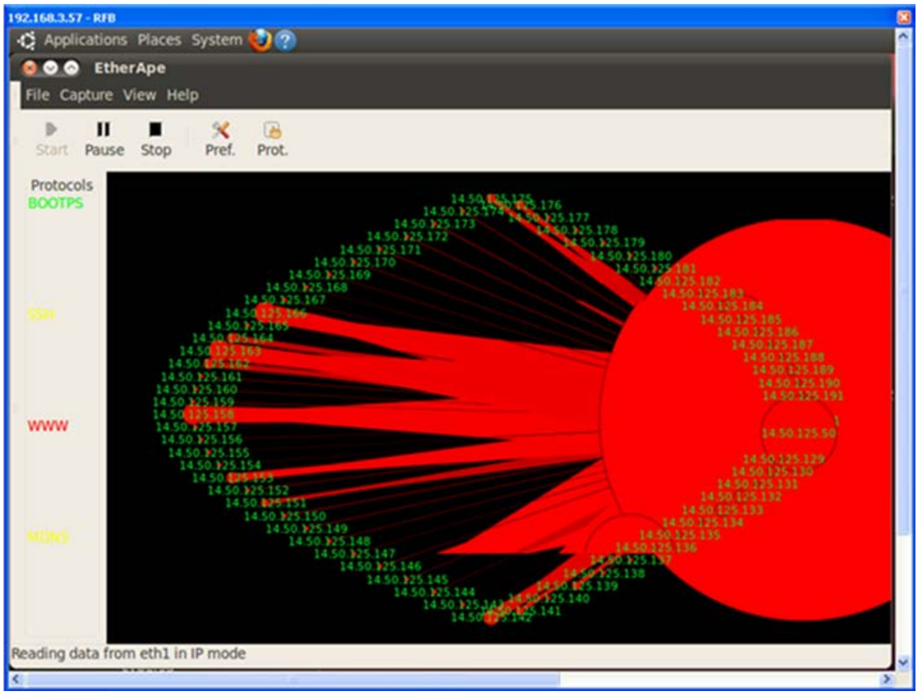
Select an interface to view input and output Kbps

🟩 UDP: 0 🟨 TCP: 16 🟥 Total: 16

🟨 Input Kbps: 390 🟥 Output Kbps: 666

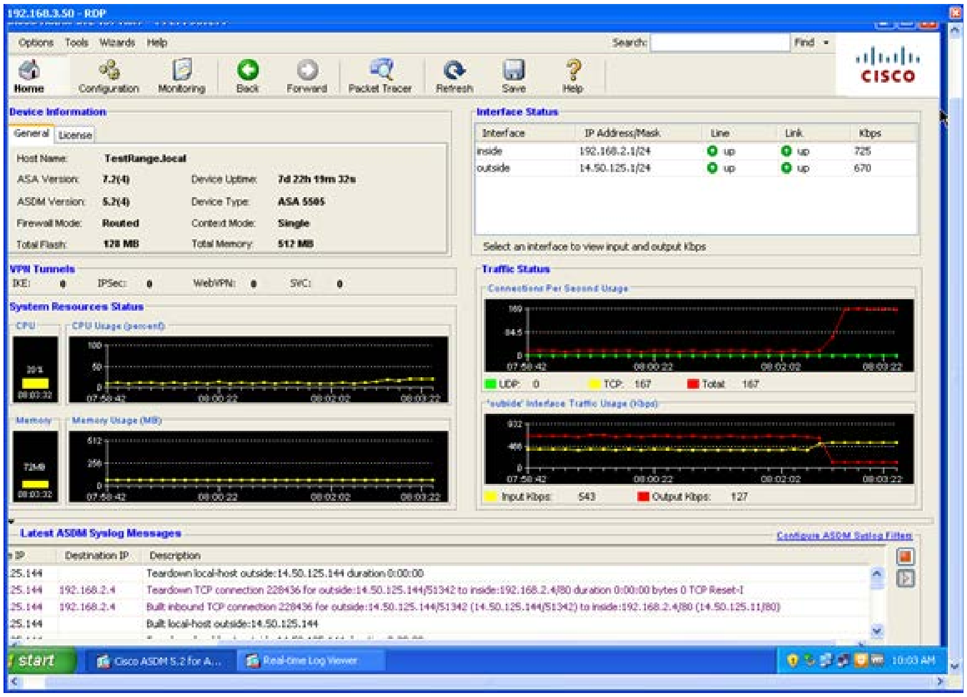
```
e IP      Destination IP  Description
125.187   192.168.2.4     Teardown TCP connection 222139 for outside:14.50.125.187/45045 to inside:192.168.2.4/80 duration 0:00:00 bytes 3868 TCP FINs
125.142                   Teardown local-host outside:14.50.125.142 duration 0:00:00
125.142   192.168.2.4     Teardown TCP connection 222138 for outside:14.50.125.142/45044 to inside:192.168.2.4/80 duration 0:00:00 bytes 4151 TCP FINs
125.187   192.168.2.4     Built inbound TCP connection 222139 for outside:14.50.125.187/45045 (14.50.125.187/45045) to inside:192.168.2.4/80 (14.50.125.11/80)
125.187                   Built local-host outside:14.50.125.187
```

Figure 6. Cisco ASDM network overview showing that the network conditions are clear. The top right plot indicates the number of active connections; the bottom right plot provides traffic volumes.



Reading data from eth1 in IP mode

Figure 7. EtherApe network visualization tool displaying normal network communications. Rays indicate communicating systems. The ray width indicates data rates.



Device Information

Host Name: TestRange.local
ASA Version: 7.2(4) Device Uptime: 7d 22h 19m 32s
ASDM Version: 5.2(4) Device Type: ASA 5505
Firewall Mode: Routed Context Mode: Single
Total Flash: 128 MB Total Memory: 512 MB

VPN Tunnels

IKE: 0 IPSec: 0 WebVPN: 0 SVC: 0

System Resources Status

Interface Status

Interface	IP Address/Mask	Line	Link	Kbps
inside	192.168.2.1/24	● up	● up	725
outside	14.50.125.1/24	● up	● up	670

Select an interface to view input and output Kbps

🟩 UDP: 0 🟨 TCP: 167 🟥 Total: 167

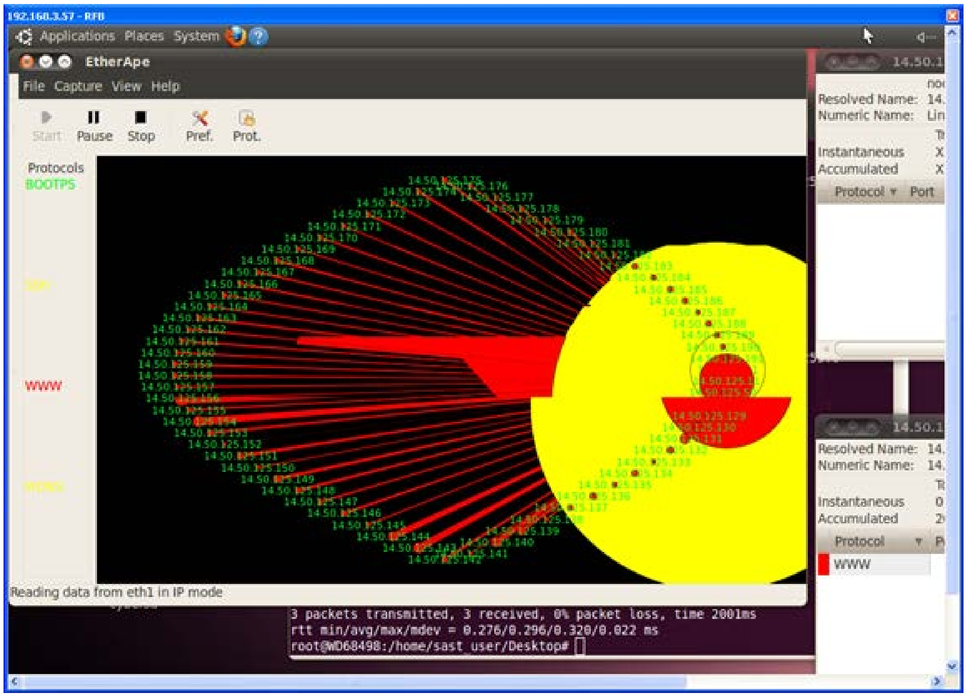
🟨 Input Kbps: 543 🟥 Output Kbps: 127

Latest ASDM Syslog Messages

Configure ASDM Syslog Filters

```
e IP     Destination IP  Description
25.144                   Teardown local-host outside:14.50.125.144 duration 0:00:00
25.144   192.168.2.4     Teardown TCP connection 228436 for outside:14.50.125.144/51342 to inside:192.168.2.4/80 duration 0:00:00 bytes 0 TCP Reset-I
25.144   192.168.2.4     Built inbound TCP connection 228436 for outside:14.50.125.144/51342 (14.50.125.144/51342) to inside:192.168.2.4/80 (14.50.125.11/80)
25.144                   Built local-host outside:14.50.125.144
```

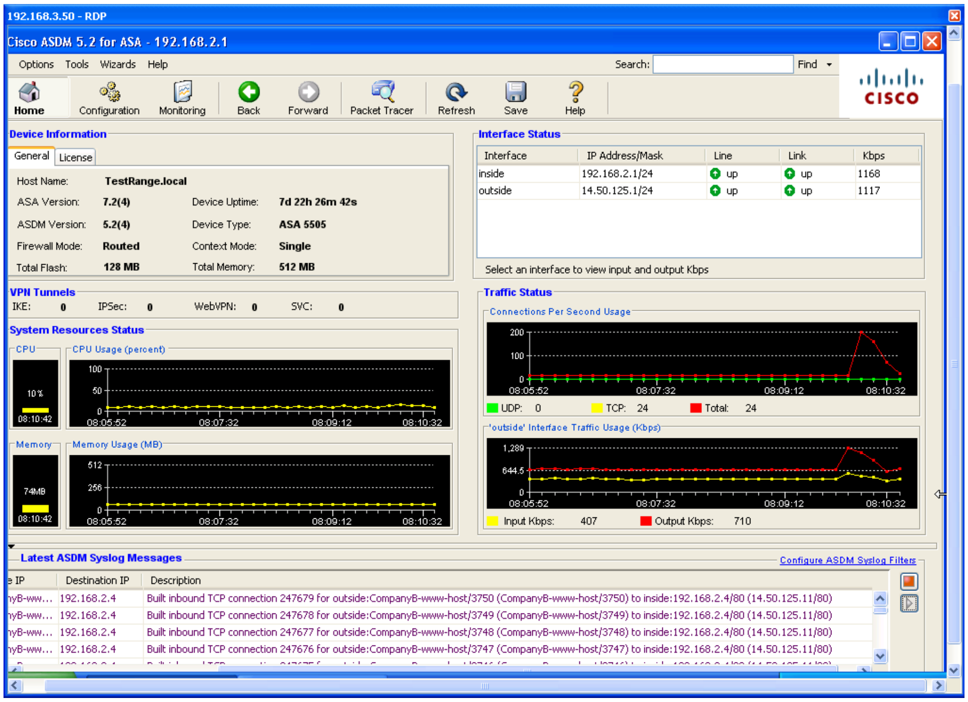
Figure 8. Cisco ASDM network overview showing that the network conditions are clear. The top right plot indicates the number of active connections; the bottom right plot provides traffic volumes.



```
3 packets transmitted, 3 received, 0% packet loss, time 2001ms
rtt min/avg/max/mdev = 0.276/0.296/0.320/0.022 ms
root@WD68498:/home/sast_user/Desktop#
```

Figure 9. The EtherApe visualizing network conditions of Paper Clip Inc.'s failed WWW service. The number of rays increase and the rays are thinner. Furthermore, SSH traffic, used to control devices, dominates the amount of traffic.

We also employed the Low Orbit Ion Cannon (LOIC), an open source network stress testing tool, to generate a Denial-Of-Service (DoS) attack against the order processing application. The application would continue to function, but the performance degradation slows transactions, and, thus, harms customer experience. During the attack, the ASDM display indicated high data rates and higher than expected number of connections. EtherApe displayed the attack as a very broad link between the WWW server and the system that is commencing the attack. Big Brother would alarm that the database server was experiencing heavy CPU utilization. Figure 10 and Figure 11show how the attack is reflected in ASDM and EtherApe.



Device Information

Host Name: TestRange.local
ASA Version: 7.2(4) Device Uptime: 7d 22h 26m 42s
ASDM Version: 5.2(4) Device Type: ASA 5505
Firewall Mode: Routed Context Mode: Single
Total Flash: 128 MB Total Memory: 512 MB

Interface Status

Interface	IP Address/Mask	Line	Link	Kbps
inside	192.168.2.1/24	up	up	1168
outside	14.50.125.1/24	up	up	1117

Select an interface to view input and output Kbps

VPN Tunnels
IKE: 0 IPSec: 0 WebVPN: 0 SVC: 0

System Resources Status

Traffic Status

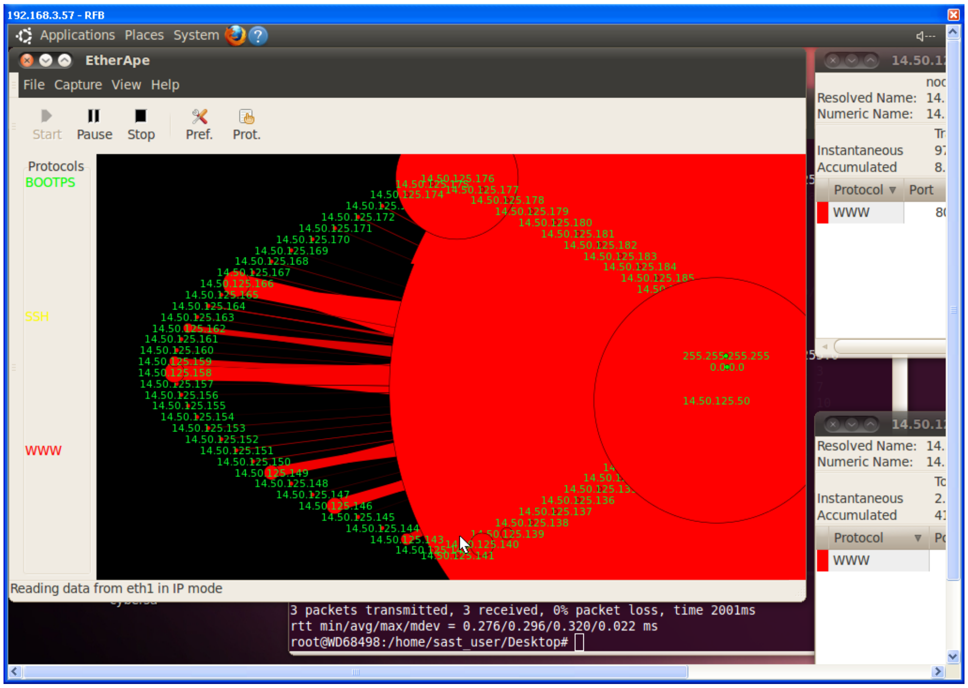
🟩 UDP: 0 🟨 TCP: 24 🟥 Total: 24

🟨 Input Kbps: 407 🟥 Output Kbps: 710

Latest ASDM Syslog Messages

e IP	Destination IP	Description
nyB-ww...	192.168.2.4	Built inbound TCP connection 247679 for outside:CompanyB-www-host/3750 (CompanyB-www-host/3750) to inside:192.168.2.4/80 (14.50.125.11/80)
nyB-ww...	192.168.2.4	Built inbound TCP connection 247678 for outside:CompanyB-www-host/3749 (CompanyB-www-host/3749) to inside:192.168.2.4/80 (14.50.125.11/80)
nyB-ww...	192.168.2.4	Built inbound TCP connection 247677 for outside:CompanyB-www-host/3748 (CompanyB-www-host/3748) to inside:192.168.2.4/80 (14.50.125.11/80)
nyB-ww...	192.168.2.4	Built inbound TCP connection 247676 for outside:CompanyB-www-host/3747 (CompanyB-www-host/3747) to inside:192.168.2.4/80 (14.50.125.11/80)

Figure 10. The ASDM network overview displaying the DoS attack. Both network connections and traffic volumes increase and are volatile.



Reading data from eth1 in IP mode

```
3 packets transmitted, 3 received, 0% packet loss, time 2001ms
rtt min/avg/max/mdev = 0.276/0.296/0.320/0.022 ms
root@WD68498:/home/sast_user/Desktop#
```

Figure 11. EtherApe visualizing the DoS attack. The large circles inform the administrator of the source and target of the attack.

5.2 Method

5.2.1 Participants

Participants were four PNNL network operations staff, solicited via email as study volunteers. The invitation stated that they were invited to participate in several simulated scenarios as part of a study on network monitoring and security. The participants originate from different parts of PNNL and perform different job functions. Summarizing their jobs, two participants provide IT support for cyber security and national security research and development groups, another participant provides IT support for PNNL's various business applications, and the last preforms a variety of tasks in support of general scientific computing. While every participant understood the concepts and skills necessary in performing the tasks at hand, only one (Participant 4) had previous experience in operating in a small business climate as illustrated by the pilot study scenario. Two participants (Participant 1 and Participant 2) have an intermediate level of experience and one (Participant 3) has little experience. All participants were familiar with the use of Big Brother network/system monitoring tool; in fact, all use it daily as part of their jobs. The experienced participant had some minimal level of exposure to the Cisco ASA and its ASDM overview page. No other participant had any prior experience. Finally, none of the participants had any exposure to EtherApe.

5.2.2 Procedure

Participants attended an initial orientation session, including all participants at the same time, lasting approximately an hour. Information about the study was provided in a consent form that each participant read and signed. They were told that they were each going to play the role of an individual system administrator/network security employee for a small company involved in Internet commerce. The network topology was reviewed and each participant was given about ten minutes of general familiarization instructions, as well as a discussion about possible cyber security exploits and attacks (including DDOS) and the company's priorities involving the need to maintain their online sales operation. At the conclusion of the orientation session, each participant signed up for an individual session that would last two hours. These sessions all occurred within a week of the orientation session.

The two-hour experimental session was divided into two main parts: (a) a one-hour portion for hands on training and familiarity with the simulation environment and the network traffic/system administration responsibilities in the scenario; (b) a one-hour test session. In (a) we introduced the operation of test bed's monitoring tools and showed how various network and system conditions were reflected in each. While all of the participants where conceptually familiar with network firewall operations, none of them had applied experience in which they operated the devices as a function of their jobs. Thus, we exerted great effort instructing the participants on the operation of the test bed's firewall: we showed them how to insert, disable, and reorder rules, how to monitor network flows, and how to capture and investigate packet traces. Part (b) was subdivided into two phases. In the first phase, we caused the termination of the ordering processing application. To remediate the fault, the participants must restart the WWW service. The scenario progressed such that the participants encountered and remediated the fault three times. In the second phase, we activated a DoS attack against the WWW service. At first glance, the attack caused symptoms similar to an unexpected application termination. Restarting the WWW service would not rectify the problem. An effective response would be to author Firewall rules to block the offending traffic and to insert the rules into the firewall.

All participants were furnished with BB and Cisco ASDM monitoring tools. Two participants (Participant 2 and Participant 4, randomly selected) were additionally furnished with EtherApe, which represented the "enhanced visualization" condition.

5.3 Results

To review, all participants were furnished with BB and Cisco ASDM monitoring tools, and two were also provided with EtherApe as an additional visualization aid.

Participant 1 (without EtherApe) – From our perch, it was not evident if the participant was choosing to use either Big Brother or the ASDM panel. During the first phase of the exercise, he relied on the alerts provided by the help desk before remediating problems. Due to a technical difficulty, the attack never reached a point to harm the ordering system. He did note the attack on the ASDM general overview panel and choose to ignore it.

Participant 2 (with EtherApe) – This participant was hyper vigilant, but unfortunately, choose to commit cyber friendly fire. He relied heavily on the ASDM overview display and noted problems nearly instantaneously. During the first phase, he reacted on the information by disabling the external interface of the ASA firewall device—in effect he choose to cut his network off from the Internet, thus committing textbook definition of cyber friendly fire. After we witnessed the participant act in this way twice, we informed the participant of the consequences. Unfortunately this was to no avail, as the participant disabled the external interface and was attempting to disable the internal interface during the DoS attack. If he was successfully, he would no longer have remote access to administrate the firewall.

Participant 3 (without EtherApe) – He was the least experienced of the four participants. He preferred ASDM overview display over the Big Brother status display. He was methodical. In the first phase, he gathered available information before deciding on the course of restarting the ordering system. The deliberate approach was slow. While the preceding participants took under two minutes to note and correct the problem, it took him at least two minutes before taking corrective action. Four minutes passed before he realized the advent of the DoS attack and it took another three minutes before he decided to take any action.

Participant 4 (with Etherape) – He was the most experienced of all the participants. While not part of his daily job function, he has served as a network operator in a part time job. During the exercise he was leaning back in his chair watching the EtherApe visualization or hunched over staring at the ASDM overview display. He response was rapid and in most cases remediated problems in under thirty seconds. Not once did we need to announce the occurrence of an event. He noted the DoS attack immediately from both the EtherApe and ASDM displays. He performed a packet trace and identified the source as AZ Payment Processor and that the attacking system was a critical component in processing payment transactions. He recognized that making changes to the firewall may cause harm later on; he would prefer contacting AZ Payment Processor first.

5.4 Discussion

It is not possible to draw generalizations from the small number of participants, particularly because of technical difficulties that affected performance of Participant 1 (and perhaps also because of existence of doubts about whether or not Participant 2 understood the instructions sufficiently to follow directions). Focus on results obtained for Participants 2 and 3 yields precious little data upon which to draw conclusions.

At a shallow level of analysis, we note that Participant 4 (who received the enhanced visualization condition) performed much better than Participant 3 (who did not receive the enhanced visualization condition). Besides the obvious conclusion that the experimental manipulation was effective, there are other possible explanations due to uncontrolled confounding factors: For example, Participant 4 had more experience than Participant 3. Because of time and budget constraints, we were unable to conduct a somewhat larger pilot study that could incorporate appropriate controls (such as a pretest-post-test design).

While we did not identify objectives related to training, some observations from the pilot study suggest training implications. Even when informed of the consequences of his choices, Participant 2 continued to engage in actions that resulted in cyber FF. This could have been a result of lack of experience with network and firewall operations, or possibly he missed the "message" communicated during the orientation session about the importance of maintaining business operations; or perhaps he believed that he was, in fact, taking the best course of action. In any case, this observation suggests that training approaches should be considered.

Because this was a pilot study, limitations and difficulties were not unexpected. Nevertheless, we still may conclude that the results at least suggest that one can demonstrate cyber FF performance differences that possibly can be related to the independent variable studied (visualization support); and perhaps more importantly for the present purposes, the Unclassified Cyber Test Range that we utilized at PNNL appears to be capable of supporting experimental studies of cyber FF. This point is important going forward, since it reinforces the recommended research strategy of conducting more controlled scientific studies of cyber SA and cycler FF in a high fidelity simulated environment. The following considerations inform the design of future experimental studies:

- Access to a larger pool of participants is needed to allow for the possibility of statistically significant results

- Participants should have a more relevant background and experience with the type of enterprise and network represented in the scenario. Participants should be fluent in the technical skills required to perform necessary actions. Experience ought to be controlled as a factor in the study.

It is important to note that the PNNL Cyber Test Range is configured to enable experiments to be conducted with remotely located participants. This means that there is potential for much wider access to participants; they can connect via the SEAL client to the Test Range from anywhere in the world, thus removing the need for the participants to be travel to PNNL's campus.

Much more specific recommendations about the design and human factors of the Cyber Test Range were derived from observing the participants obtaining feedback from the participants:

- While the client does ease the consolidation of several displays onto one workstation, the participants are limited by the amount of screen real-estate offered by that workstation. The participants were furnished with two 19" widescreen LCD monitors and a large, overhead projected display that was approximately 60". Most of the time, the participants located the ASDM general overview on one monitor and located either Big Brother or EtherApe on the overhead display (the other monitor was used to perform tasks on the systems). Perhaps additional display options (more large 21+ inch monitors) should be provided.
- The overhead display was aimed too high: participants either leaned back in their chairs or cranes their necks to view it.
- Within SEAL, the TAB, CTRL +, ALT +, and META + key sequences are not passed to the remote session, but are interpreted locally on the workstation. For example, the key sequence CTRL + S is normally assigned the function of saving changes to a file. When using SEAL, the participants had to use the mouse and activate the save via the application's file menu.
- Two participants voiced frustration with the inability to label windows. SEAL identifies windows by network address and these participants stated that they prefer identifying the windows with more descriptive labels, such as the system's purpose.
- Finally, all the participants voiced frustration with the latency of screen updates. Screen updates can be slow and graphical elements such as plots do not quickly refresh themselves under the client.

In summary, the Unclassified Cyber Test Range is an able resource in studying cyber FF. Future studies will require a greater number of participants with similar experience and who can demonstrate the skills required by the study scenarios. More time should be allocated for the participants to become familiar with the scenario environment. Furthermore, participants should have access to a greater number of large (21" or greater) monitors.

[This Page Intentionally Blank]

6. Conclusions and Future Direction

Research in cyber FF should be founded upon scientific principles and empirical studies in human factors and cognitive engineering, such as seminal human factors work on SA by Endsley (1995) and later by Tada and Salerno (2010), who mapped constructs of SA to more cyber-relevant network environments. The research performed by PNNL on cyber FF has helped to define research questions and has laid a foundation for the construction of a cyber SA/FF test range that can provide a framework for empirical investigations of factors contributing to the cyber FF phenomenon and impacts on performance of proposed mitigations that can be in the form of training/awareness or decision aids.

An advanced concept that is being pursued by PNNL cyber security programs is the notion of Asymmetric Resilient Computing (ARC), which has been characterized by goals to set up cyber infrastructure and network architectures that present a “moving target” to potential attackers in an attempt to overcome and hopefully reverse the current asymmetric state of cyber war that favors the adversary. The goals and challenges of this program align with issues that we have articulated in our current cyber FF research. This research also directly meets essential needs of DOE cyber security and counterintelligence (DOE CIO, DOE-IN) as well as cyber security programs within the DoD and the intelligence community.

The fundamental research goal is to develop a scientific understanding of the behavioral implications of cyber FF. Research is needed to extend our current understanding of cyber SA and to develop metrics and measures for cyber FF. The principal scientific research questions include: What are root causes of cyber FF? What are possible mitigating solutions, both human factors and technical/automated? We have examined relevant research and cognitive theory, and we have begun to establish laboratory assets with which to conduct empirical research to address these questions through the development of appropriate test scenarios. Now, follow on research is required to investigate the phenomenon and relevant contributing factors as well as mitigation strategies, utilizing experimental methods that may be conducted within facilities such as the cyber SA/FF test bed at PNNL. A major objective of this research is investigation of approaches to and assessment of effectiveness of cyber FF mitigation strategies, such as training and decision aids/tools.

The proposed research will help to advance the general field of cyber SA—research and theory produced in this effort will inform other ongoing cyber security research. In addition, the proposed research will facilitate the design and prototyping of automated or semi-automated systems (or decision aids) to help increase cyber SA and eliminate or decrease cyber FF; this provides a foundation for development of commercial products that enhance system effectiveness and resiliency.

[This Page Intentionally Blank]

7. References

Allnut, M. (1982). Human factors: Basic principles. In R. Hurst, & L. R. Hurst (Eds.), *Pilot error* (pp. 1–22). New York: Aronson.

Andrews, D. H. and K. T. Jabbour (2011). "Mitigating Cyber Friendly Fire: A Sub-Category of Cyber Mishaps." High Frontier 7(3):5-8

Barthol, R. P., & Ku, N. D. (1959). Regression under stress to first learned behavior. *Journal of Abnormal and Social Psychology, 59,* 134–136.

Davies, D. R., & Tune, G. S. (1970). *Human vigilance performance*. London: Staples Press.

Driskell, J. E., Salas, E., & Johnston, J. (1999). Does stress lead to a loss of team perspective? *Group dynamics: Theory, research and practice, 3,* 291–302.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 1873–201.

Endsley, M. R. (1995). "Towards a Theory of Situation Awareness in Dynamic Systems." Human Factors **37**(1): 32-64.

Endsley, M. R. and W. M. Jones (1997). Situation Awareness, Information Dominance, and Information Warfare. Wright-Patterson AFB, OH, United States Air Force Armstrong Laboratory.

Endsley, M. R. and W. M. Jones (2001). A Model of Inter- and Intrateam Situation Awareness: Implications for Design. New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments. M. McNeese, E. Salas and M. Endsley. Santa Monica, CA, Human Factors and Ergonomics Society.

Friedman, I. A., & Mann, L. (1993). Coping Patterns in Adolescent Decision-Making: An Israeli-Australian Comparison, *Journal of Adolescence, 16,* 187-199.

Greitzer, F. L., & DH Andrews. (2010). Training strategies to mitigate expectancy-induced response bias in combat identification: A research agenda. Chapter 11 in D.H. Andrews, R.P. Herz, and M. B. Wolf (Eds), *Human factors issues in combat identification*, Don Harris, Neville Stanton, and Eduardo Salas (Series Eds.) *Human factors in defence*. Surrey: Ashgate. 173-190. PNNL-SA-64673

Greitzer, F. L., S. L. Clements, et al. (2009). Towards a Research Agenda for Cyber Friendly Fire, Pacific Northwest National Laboratory.

Heuer, Richards J., Jr, (1999). "Chapter 8: Analysis of Competing Hypotheses", *Psychology of Intelligence Analysis*, Center for the Study of Intelligence, Central Intelligence Agency.

Hollan, J., Hutchins, E., and Kirsh, D.(2000). Distributed Cognition: Toward a New Foundation for Human-Computer Interaction Research. *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 2, June 2000, Pages 174–196.

Janis, I. L., & Mann, L. (1977). *Decision Making*, New York: The Free Press. Kavanagh, J. (2005). *Stress and performance: A review of the literature and its applicability to the military.* (RAND TR-192, ADA439046). Santa Monica CA: Rand Corp. [Online]. Available at http://www.rand.org/pubs/technical_reports/2005/RAND_TR192.pdf [accessed: February 12, 2008].

Jones, M.D. (1998). *The Thinker's Toolkit: Fourteen Powerful Techniques for Problem Solving*. New York: Three Rivers Press.

Klein, G. (1996). The Effects of Acute Stressors on Decision-Making. In J. Driskell, & E. Salas (Eds.), *Stress and human performance* (pp. 49-88). Hillsdale, NJ: Lawrence Erlbaum.

Kohn, H. (1954). Effects of variations of intensity of experimentally induced stress situations upon certain aspects of perception and performance. *Journal of Genetic Psychology, 85*, 289–304.

Kozlowski, S. W. J. (1998). *Training and Developing Adaptive Teams: Theory, Principles, and Research.* Making Decisions Under Stress: Implications for Individual and Team Training. J. A. Cannon-Bowers and E. Salas. Washington DC, America Psychological Association.

Leavitt, J. (1979). Cognitive demands of skating and stick handling in ice hockey. *Canadian Journal of Applied Sport Sciences, 4,* 46–55.

Norman, D. A. (1988). The psychology of everyday things.

Sarter, N. and D. Woods (1991). "Situation Awareness: A Critical But Ill-Defined Phenomenom." International Journal of Aviation Psychology **1**(1): 45-57.

Smith, M. D., & Chamberlin, C. J. (1992). Effect of adding cognitively demanding tasks on soccer skill performance. *Perceptual and Motor Skills, 75*, 955–961.

Staal, M. A. (2004). *Stress, cognition, and human performance: A literature review and conceptual framework.* (NASA/TM-2004-212824). Moffett Field, CA: NASA Ames Research Center.

Tadda, G.P., and Salerno, J.S. (2010). "Overview of cyber situation awareness." In S. Jajodia et al. (eds.), *Cyber Situational Awareness: Issues and Research*. New York: Springer, 15-35. http://www.springer-link.com

United States Air Force Chief Scientist (AF/ST) (2010). Report on Technology Horizons A Vision for Airforce Science & Technology During 2010-2013.

van Overschelde, J. P., & Healy, A. F. (2001). Learning of nondomain facts in high- and low-knowledge domains. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1160-1171.

Zajonc, R. B. (1965). Social facilitation. *Science, 149*, 269–274.