Security Informatics
a SpringerOpen Journal

RESEARCH Open Access

Security informatics research challenges for mitigating cyber friendly fire

Thomas E Carroll[1*], Frank L Greitzer[2] and Adam D Roberts[1]

Abstract

This paper addresses cognitive implications and research needs surrounding the problem of cyber friendly fire (FF). We define 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. We describe examples of cyber FF and discuss how it fits within a general conceptual framework for cyber security failures. Because it involves human failure, cyber FF may be considered to belong to a sub-class of cyber security failures characterized as unintentional insider threats. Cyber FF is closely related to combat friendly fire in that maintaining situation awareness (SA) is paramount to avoiding unintended consequences. 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 its nodes, the nature of the activities or work performed, and the available defensive and offensive countermeasures that may be applied to thwart network attacks. We describe a test bed designed to support empirical research on factors affecting cyber FF. Finally, we discuss mitigation strategies to combat cyber FF, including both training concepts and suggestions for decision aids and visualization approaches.

Introduction

Computer and network security are among the greatest challenges to maintaining effective information systems in public, private, and military organizations. In defining computer security, Landwehr [1] described three threats to Information systems: (1) the unauthorized disclosure of information, (2) the unauthorized modification of information, and (3) the unauthorized withholding of information (e.g., denial of service or DoS). Denning [2] referred to information warfare as a struggle between an offensive and a defensive player over an information resource, with outcomes that may affect availability or integrity of the resource. While much attention has been devoted to combating external threats such as worms, viruses, and DoS attacks, actions by insiders pose a significant threat to computer and network security. This insider threat, however, is not confined to "bad actors" that intentionally perform malicious acts against an information system. Just as unintended actions by friendly forces may impact physical resources and security of friendly forces in military engagements, the actions of well-intentioned cyber defenders may result in harm to information resources and security. The focus of the present paper is on understanding and mitigating threats to intelligence and security informatics posed by cyber friendly fire.

While friendly fire (FF) is a familiar term, cyber FF is a relatively new concept for the information security community. An initial proposed definition of cyber FF, from Greitzer et al. [3], emphasizes three key characteristics:

- Cyber/electronic actions are performed intentionally,
- Actions are offensive or defensive,
- Actions result in inhibiting, damaging, or destroying friendly or neutral infrastructure or operations.

Andrews and Jabbour [4] 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.

*Correspondence: Thomas.Carroll@pnnl.gov
[1]Pacific Northwest National Laboratory, P.O. Box 999, 99352 Richland, Washington, USA
Full list of author information is available at the end of the article

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. 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" [5]. 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 [6]. 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.

Given these considerations, a revised definition of cyber FF [6] is:

> *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.** As a cost saving measure, Company XYZ has outsourced their corporate website and email to a hosting company. A hacker who has compromised and is now on the hosting company's infrastructure disrupts services by attempting to break into Company XYZ's resources. An administrator at Company XYZ notes the activity and quickly takes actions to protect company resources by blocking traffic from network addresses that are the source of the attack. As a direct consequence of these actions, Company XYZ employees lose access to their corporate website and email.
>
> **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 XYZ, who relies on the software as a critical component of their e-business platform, rapidly deploys the fix on their infrastructure. The patch introduces a problem into the software, causing transactions to fail and frustrating potential customers who are attempting to purchase the company's products.

The next examples illustrate defensive actions that unintentionally 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 sensitive 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 hastily writes a new firewall rule to block suspected malicious network traffic. He errors in composing the rule, but before he catches his mistake, he publishes the errant rule to production. The rule disrupts the operations of the company's web servers, which inhibits purchases, harming sales.

Cognitive approaches to cyber friendly fire research

The concept of cyber FF is similar in many respects to combat friendly fire [3], and like combat friendly fire, a fundamental cognitive issue lies in maintaining situation awareness (SA). In addition, cyber FF is closely related to some aspects of insider threat, especially when viewed within the broad framework of cyber security failures. This section provides some background and perspective on the cognitive foundations for cyber FF.

Cyber security failures and the unintentional insider threat

The domain of cyber security spans a broad spectrum of research and operational policies to address outsider cyber threats, insider threats, and other failures such as accidents or mishaps. Figure 1 provides a conceptual view of where cyber FF fits within this broader framework of cyber security failures.

Included in the framework is the familiar branch (shown in the black boxes) representing cyber attacks and exploits by malicious insiders, the latter representing the most highly studied insider threat research

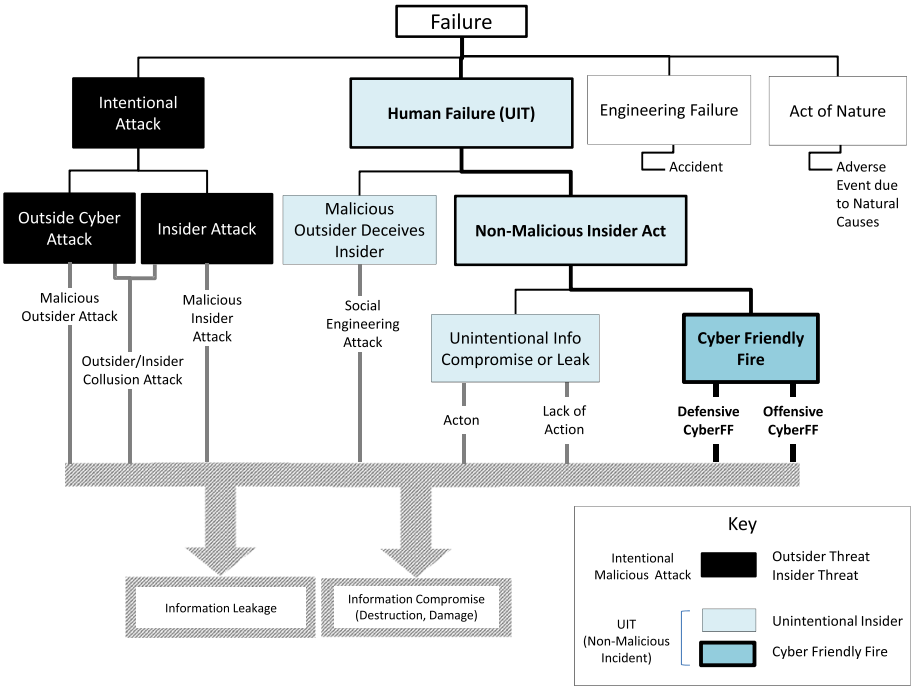

Figure 1 Conceptual framework of cyber security failures. The embolden path illustrates that cyber FF incidents is a subclass of unintentional insider threat incidents.

topic. Engineering failures that may be attributed to system/hardware/software vulnerabilities are represented in unfilled boxes. Of most interest for the present discussion are the branches of the hierarchy that relate to human failures that may be attributed to actions of insiders. The topic of unintentional insider threat (UIT) has been largely ignored until recent research by [7] that has provided a working definition of UIT:

> An unintentional insider threat is (1) a current or former employee, contractor, or business partner (2) who has or had authorized access to an organization's network, system, or data and who, (3) through action or inaction without malicious intent, (4) unwittingly causes harm or substantially increases the probability of future serious harm to the confidentiality, integrity, or availability of the organization's resources or assets, including information, information systems, or financial systems

As pointed out in [7], UIT incidents share a common characteristic in which an organizational insider facilitates the actual or potential threat event. However, there is a distinction between the UIT cases that originate with actions performed by the internal, non-malicious member of the organization, versus UIT events that originate with an outside malicious agent (outside agent may recruit a malicious insider to participate in a collusion attack, or outside agent may deceive a non-malicious insider to take actions that enable an attack). These various cases are depicted in separate branches of Figure 1. The research reported in [7] relates primarily to the light blue boxes in Figure 1, especially social engineering exploits, information leakage and information compromise. For these UIT cases, there are four main types of incidents, which are referred to in [7] as UIT threat vectors: (a) accidental disclosure (DISC), (b) attack enabled through use of malicious code such as malware or spyware (UIT-HACK), (c) improper disposal of physical records (PHYS), and (d) lost or stolen portable equipment (PORT).

We suggest that the cyber FF definition clearly fits within the above broad definition of UIT. However, UIT research to date has not considered the case of cyber friendly fire; for example, cyber friendly fire is not mentioned in [7] nor are any cases included in their discussion or taxonomic descriptions. This is not surprising since there is no repository of such cases to draw from. By the same token, our original work on cyber FF emphasized the differences between cyber FF and insider threat, maintaining that they are distinct cyber threats. This is true if one considers only malicious insider threats, but as we now argue, cyber FF should legitimately be considered as a special case of UIT.

Therefore, it is useful to examine UIT research and associated UIT mitigation strategies to identify possible

approaches for addressing cyber FF. The area of greatest commonality between UIT and cyber FF is in human performance failures. As noted in [7]: "A major part of the UIT definition is the failure in human performance. While human errors can never be eliminated completely, they can be dramatically reduced through human error mitigation techniques. Such techniques should focus on system conditions that contributed to…the resulting errors and adverse outcomes." These remarks and associated suggestions for enhancing the decision maker's situation awareness and reducing human errors pertain just as strongly to cyber FF mitigation strategies as they do to UIT threat mitigation. For this reason, we may use the arguments and suggestions provided in [7] to provide high level organizational and human factors strategies for combating cyber FF.

Problems associated with organizational factors, such as work setting, management systems, and work planning, may impact employee performance. For example, job stress [8] and time pressure [9] negatively affect performance; heavy and prolonged workload can cause fatigue, which adversely affects performance [10]. Moreover, organizational factors that increase stress may in turn lead to human factors/cognitive impacts such as narrowing of attention (attending to fewer cues) [11,12] and reduced working memory capacity [13-15]. Cognitive factors associated with UIT susceptibility include attention deficits and poor situation awareness [16,17], lack of knowledge and memory failures [18-20], and high workload or stress that impairs performance or judgment [10,21]. Finally, external or organizational factors may affect an individual's emotional states, both normal and abnormal, which in turn can affect the human error rate and lead to UIT occurrences.

Cognitive systems perspective

The traditional approach in accounting for performance failures such as combat friendly fire/fratricide and the lesser-examined cyber FF is to regard these events as aberrations—failures of an individual or a system. As with most performance failures (errors), assigning blame to the individual(s) responsible for a cyber FF incident is not a sufficient mitigation strategy: there is typically no single cause of these errors that occur in the "fog of war". To understand the causes (and persistence) of cyber FF, it is necessary to consider the human factors, and it seems particularly relevant to address the problem from a cognitive systems/naturalistic decision making perspective (e.g., [22,23]). Thus, we should ask: *How did the individual perceive the situation? Why did the individual see the event that way? Why did the individual act in a way that turned out to be erroneous?*

A cognitive systems perspective leads us to consider research on SA and mental models. The SA scientific literature 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 [16]: 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). Later work by McGuiness and Foy [24] added the *resolution* of the situation (Level 4 SA), which is deciding on a single course of action from a set of possible actions to achieve the required outcome to the situation.

SA depends on an accurate mental model [25]. Mental models have been described as well-defined, highly organized, and dynamic knowledge structures that are developed over time from experience (e.g., [26]). 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 [27].

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?

Tadda and Salerno [28] 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 Tadda 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"[a]. 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 percent 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 units*) are required to invoke this anticipatory process:

- Knowledge of each enterprise, enterprise's network structure, and network component
- 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 [16] and in particular, the extent to which team members possess the same SA on *shared* SA requirements [29,30]. 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.

Given these considerations, a recommended approach is to capture the mental models that constitute the above types of knowledge, and then to tailor training approaches and tools to address associated cognitive factors.

Trends that make digital SA harder

Current trends challenge the abilities of individuals and teams to maintain digital SA; more particularly, changes in the roles and communications among cyber security professionals, as well as paradigmatic changes brought about by cloud and utility computing, have increased the difficulty of acquiring, understanding, and maintaining critical knowledge units necessary for effective defense and operation of the information and communications infrastructure. One trend is a growing separation between the roles of individuals responsible for cyber defense mission planning and the roles of individuals responsible for operating and defending the information and communications infrastructure. Missions are defined in terms of abstract resources and quality of service attributes, rather than actual systems and devices. For example, a mission in support of a 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"). In many cases, external third parties provide and operate the infrastructure. Under these circumstances, the mission planner may not be aware of what resources are allocated, the underlying network topology, or the geographical location of the resources. The operators and defenders are compartmentalized—and often isolated—from the mission planners, understand the resources and infrastructure but are unaware of the missions that the infrastructure is serving. Communication between mission planners, operators, and defenders is complicated—if it can occur at all (e.g., "need to know" restrictions)—because planners focus on the mission, while operators and defenders focus on resources. A second trend, offered by cloud and utility computing, employs dynamic resource allocations in response to changing demands and requirements. Dynamic resource management can quickly revise and relocate allocations, as well as change the purpose or criticality of systems—this changing operational landscape challenges the ability of cyber defenders to maintain an accurate accounting of system resources, assets, and vulnerabilities. The ability to acquire and maintain critical knowledge units that are needed to support effective SA for defending these types of enterprises exceeds the efforts of an individual or even a small team—it demands the support of an entire enterprise and its complement of third parties, an extremely difficult goal. Research is needed to flesh out requirements for information sharing and for automated tools, visualization support, and decision aids to ensure that defenders

have the necessary knowledge and SA to protect their enterprise.

Mitigation approaches

In this section, we describe approaches and tools to mitigate cyber FF. Following research on organizational factors underlying human error, we first discuss organizational best practices (Section 'Organizational best practices') that are recommended to foster productive work environments, relieve stress, and reduce cognitive load. Next, in Section 'Training', we discuss relevant research on learning and cognition that may be applied to improve the effectiveness of training approaches in reducing the likelihood of cyber FF outcomes. In Section 'Effects of stress on performance', we discuss implications of this research for the design and development of tools to help enhance SA and decision making (for example, by promoting the acquisition and use of critical knowledge units elaborated in Section 'Cognitive systems perspective').

Organizational best practices

Following their discussion of possible organizational and human factors that contribute UIT, [7] described possible UIT mitigation strategies. Organizational best practices suggested in [7] that are most relevant to mitigating the incidence of cyber FF are those that help to reduce cognitive load, stress, and ultimately lead to lower risks of human errors:

- Review and improve management practices to align resources with tasks.
- Improve data flow by enhancing communication and maintaining accurate procedures.
- Maintain productive work setting by minimizing distractions.
- Implement effective work planning and control to reduce job pressure and manage time.
- Maintain employee readiness.
- Maintain staff values and attitudes that align with organizational mission and ethics.
- Implement security best practices throughout the organization.

Training

Conventional training, simulation-based training, and war gaming can each be utilized as parts of an integrated strategy to educate, raise awareness about cognitive biases and limitations, develop coping skills, and exercise skills designed to mitigate the environmental and situational factors that increase the likelihood of cyber FF. The goal of such training approaches is to provide the learner with experiences and instruction on cues, mental models, and actions that, with practice, will help establish a repertoire of well-learned concepts that can be executed under stressful or in novel, uncertain conditions. 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 [31] 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.

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 [32]. Originally observed by Kohn [33], this finding has been replicated often, first by seminal work from Easterbrook [34] demonstrating a restriction in the range of cues attended to under stress conditions (tunneling) and many other studies (see [35]). Research by Janis and Mann [36] 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 [37] 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. 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 [38].

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 [39,40], 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 [41] 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 [42-44]—even less efficient or more error prone 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.

Effects of stress on learning

Research suggests that high stress during instruction tends to degrade an individual's ability to learn. The research literature consistently demonstrates that elements of working memory are impaired, although the mechanisms behind these effects are poorly understood [35]. Stress appears to differentially affect working memory phases [45,46]. One instructional strategy to address stress effects is to use a phased approach with an initial learning phase under minimum stress, followed by gradual increasing exposure to stress more consistent with real-world conditions [31]. Similarly, 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.

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 [47] 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.

Implications

Based on the brief 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.
- 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.

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.

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 units 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" [48,49]). Thus, a decision aid that displays critical knowledge units 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 [28] 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.

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., [50]). 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 [51].

Recommendations

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
  - Develop realistic cyber war gaming scenarios that systematically vary stress (e.g., as produced by varying cognitive workload through tempo of operations and density of attacks)
  - Incorporate stress management techniques
  - 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 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.

Research test bed and preliminary studies

As concluded in Section 'Mitigation approaches', more research is needed to enhance our understanding of the factors underlying cyber FF and to explore and validate possible mitigation approaches and tools. Cognitive engineering research is needed to focus on determinants of SA deficiencies and human errors in working with tools aimed to support cyber security analyst perception and decision making processes.

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 [16] and later by Tadda and Salerno [28], who mapped constructs of SA to more cyber-relevant network environments. The present paper has sought to define research questions and to lay a foundation 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.

Along these lines, we conducted a preliminary study at PNNL to help address these research questions using the simulation capabilities of PNNL's Unclassified Security Test Range test bed. The purpose of the pilot study was to demonstrate feasibility of an experimental methodology to assess effectiveness of decision aids and visualizations for cyber security analysis. Because the experiment was limited to a very small number of participants, interpretation of results was speculative, but the design and implementation of the testbed itself serve to advance the research goals described here.

Unclassified security test range

The Unclassified Security Test Range consists of a combination of virtual and physical devices for testing, simulation, and evaluation. This closed network offers services found on a production network without the costs associated with duplicating a real environment. The idea is to duplicate enough of a real network to allow the test bed to appear realistic. In order to achieve this lofty goal, the virtual and physical environment is flexible and can be customized to represent different configurations based on requirements. With the proper configuration and orchestration of components, it is possible to create simulated environments that model Fortune 500 enterprises, and application and infrastructure service providers. The test range also has a room that is mocked up as an advanced "network operation center". Besides workstations provisioned with several large monitors, there were two large over head displays, allowing the projection of visualization such as network health and status. Observers can watch subjects from a vantage point, which is partially obscured from the participants view.

The test range creates virtual machines for user workstations and servers that interconnected using real networking switches, routers, and firewalls. Every virtual machine has at least two network interfaces, one for management and observation, and one or more for experiment network traffic. The test range features a unique simulation capability called ANTS. This software package simulates user behaviors: agents that are deployed on the virtual machines network have models or profiles of operator's use of real applications such as Microsoft Word, Outlook, and Internet Explorer. Application usage then generates the traffic found in normal networks. The advantage of this approach over others is its ability to create higher fidelity.

The test range has a network monitoring feature that provides the capability to monitor, log, and analyze all the traffic flowing through the network. Additionally, remote researchers and observers are provided a capability to view into the range.

Procedure

The test bed was configured to appear as an e-commerce website, a payment processor, and an "Internet" to ferret communication between e-commerce site and payment processor, and customers and malcontents to the e-commerce site. Participants were tasked with the role of network and security operations and were responsible for maintaining the operation of the e-commerce site. They were confronted with two types of events that interfered with customers assessing the website. 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 the order system. The second event was a Denial-of-Service (DoS) that originated from the payment processor partner. While a partner attacking appears exceptional, there have been cases in which attackers have exploited partner relationships and used compromised partners as stepping stones to further their compromise towards reaching their goals. The DoS attack consumed large quantities of resources, slowing customer access. Both events appear, at least at first glance, to be similar.

Participants were furnished with four widely available tools. The first is Big Brother (BB) system and network monitor. BB was configured to monitor various aspects of the 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 2 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 constituent systems are unable to communicate with systems on the Internet. The second tool for monitoring is the Cisco ASA's ASDM panel (shown in Figure 3). 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 was also furnished with EtherApe, a network monitor that displays network activity graphically. As depicted in Figure 4, the interface 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.

Participants

Participants were four PNNL network operations staff, solicited via email as study volunteers. The invitation

Figure 2 The Big Brother (BB) network and system monitor overview page. Each row is a network resource; each column is an indicator of a test result of the resource status. A green orb indicates healthy, while a red orb denotes a failed test.

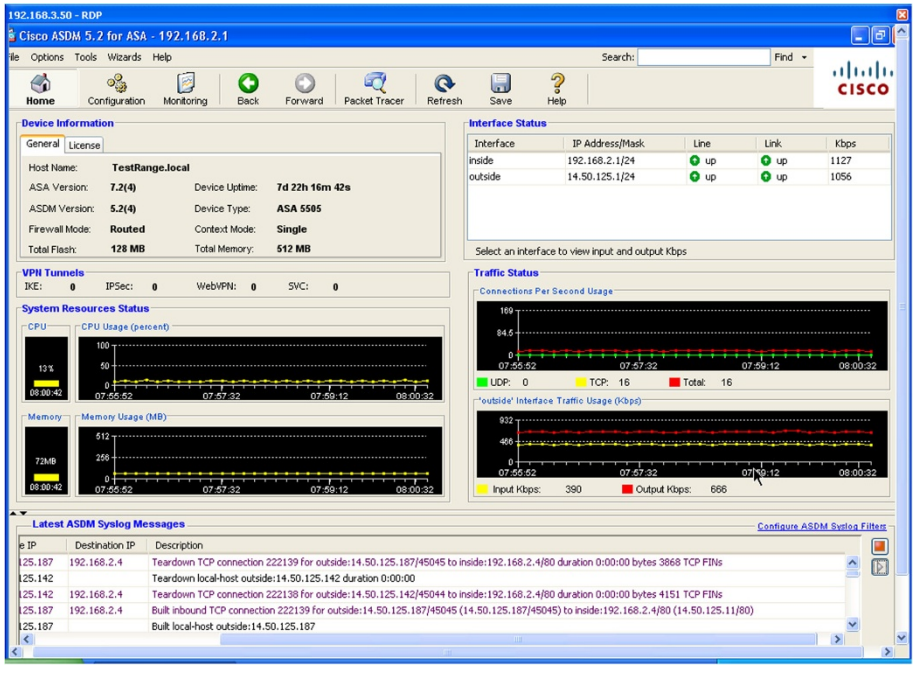

Figure 3 Cisco's ADSM network overview panel. The current status of firewall's interfaces are shown in the top right panel. The middle panels graph in real time resource utilization, connection rate, and data rate. Lastly, the bottom panel shows the device's system log messages.

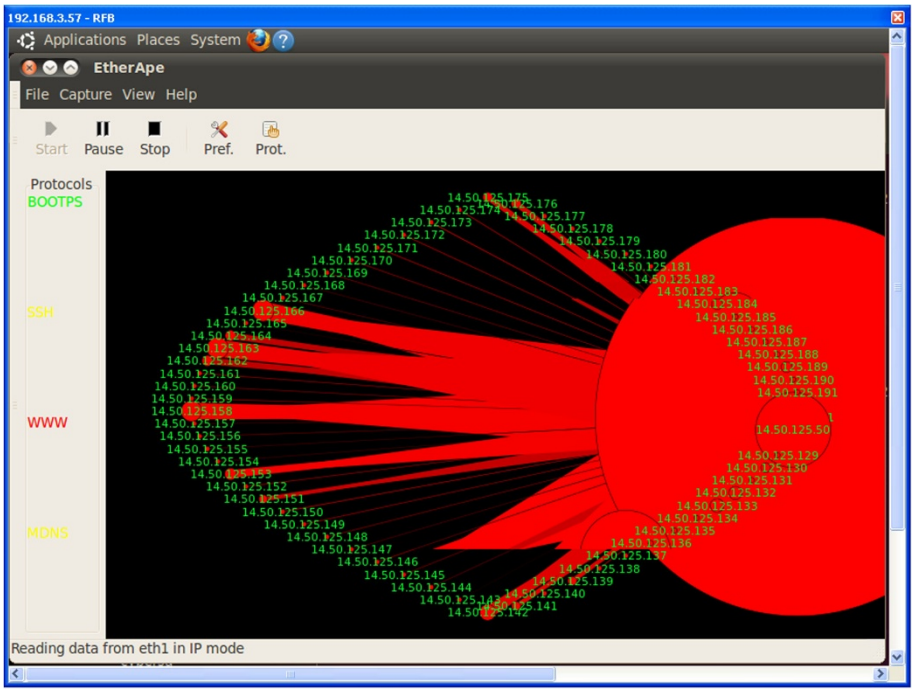

Figure 4 EtherApe network visualization tool. Rays indicate communicating systems; ray widths is proportional to data rates.

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 performs 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 had previous experience in operating in a small business climate as illustrated by the pilot study scenario. Examining the remaining three participants, two can be ranked as having intermediate level of experience and the last having 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.

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

Results

To review, all participants had ready access to BB and Cisco ASDM monitoring tools; 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 chose to ignore it.

Participant 2 (with EtherApe) This participant was hyper vigilant. Unfortunately, his choices and actions lead to cyber friendly fire. Nevertheless, his actions constitute cyber FF. He relied heavily on the ASDM overview display and noted problems nearly instantaneously. During the first phase, he reacted to the information by disabling the external interface of the ASA firewall device—in effect he chose 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 successful, he would no longer have remote access to administer the firewall.

Participant 3 (without EtherApe) 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 participants 1 and 2 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. His 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 the 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 the partner first.

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 3 and 4 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 cyber FF in a high fidelity simulated environment.

Conclusions and recommendations that serve to inform future design of such experimental studies are:

- Access to a larger pool of participants is needed to allow for the possibility of statistically significant results. The observational methods employed, interview procedures, and the performance measures collected would readily apply to an expanded study with more participants.
- 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.

Much more specific recommendations about the design and human factors of the Cyber Test Range were derived from observing the participants and obtaining feedback from the participants including deficiencies in software packages, monitor placement, choice of keyboard and mice, and the height of the overhead displays.

Conclusions and future direction

Cyber FF is one class of cyber security failures. Since it is based in human failure, we have argued that cyber FF belongs to a sub-class of cyber security failures that is characterized by unintentional insider threats. There is a tendency to regard such failures as aberrations that

can be fixed with technological advances, such as technology solutions to improve SA, to increase accuracy in identifying targets, or to improve the precision of defensive or offensive actions. We have argued that a sound mitigation strategy, whether or not incorporating technological advancements, must be focused on identifying and accommodating the realities of human performance. Contributing factors attributed to cyber FF must be empirically studied along with the benefits of training, awareness, and tools meant to reduce the number and severity of incidents. To this end, we have described a preliminary study that we conducted that demonstrates cyber FF research in a controlled, isolated testbed environment. While results where speculative due to the small number of participants, it does demonstrate that experimentation in a testbed can advance the research goals described within.

Related research is ongoing: An advanced concept that is currently being pursued by PNNL cyber security research programs is the notion of Asymmetric Resilient Cybersecurity (ARC)[b], which is characterized by goals of standing up resilient and robust 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 affairs that favors the adversary. The goals and challenges of this program align with issues that we have articulated in our research on cyber FF, particularly in ways that can be seen as amplifying the cyber FF challenge: e.g., maintaining enterprise-wide SA when the network, systems, and components "move" continuously and dynamically. Moreover, ongoing research at PNNL seeks to develop and assess visualization and decision support tools that address cognitive limitations. Current research is developing Kritikos, a network resource identification, resource dependency discovery, and criticality assessment tool. Dependencies are identified by Self-Organizing Maps (SOM), a neural network machine learning algorithm, discovering repeated spatio-temporal patterns in IP Information Flow eXport (IPFIX) record sets. Patterns in time and space indicate usage; repeated observations of a pattern suggest a dependent relationship. The patterns allow a dependency-based network model to be generated. This model is a (disconnected) graph where vertices are resources and edges indicate dependent relationships. A business process/operations model annotated with indication of network resources and business criticality, assumptions not unusual for today's enterprises, can be fused with the network model, illuminating indirect resources. Furthermore, the relationship between business process and network resources can be used to assess the criticality of resources in terms of business objectives and requirements. Cyber FF research also directly meets essential needs of DOE cyber security 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 taken some initial steps toward investigating these research questions in empirical laboratory studies using realistic test scenarios in a cyber SA/FF testbed facility [52]. Continued empirical research is required to investigate the phenomenon and relevant contributing factors as well as mitigation strategies. A major objective should be to investigate approaches to and assessment of effectiveness of cyber FF mitigation strategies, such as training and decision aids/tools. Such research promises to advance the general field of cyber SA and inform other ongoing cyber security research. In addition, it is hoped that this research will facilitate the design and prototyping of automated or semi-automated systems (or decision aids) to increase cyber SA and eliminate or decrease cyber FF; this provides a foundation for development of commercial products that enhance system effectiveness and resiliency.

Endnotes

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

[b]Information about PNNL's Asymmetric Resilient Cybersecurity (ARC) Lab Direct Research & Development Initiative can be found at: http://cybersecurity.pnnl.gov/arc.stm.

Acknowledgment

Portions of the research were also funded by PNNL's Asymmetric Resilient Cybersecurity (ARC) Laboratory Research & Development Initiative. 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.

Author details

[1]Pacific Northwest National Laboratory, P.O. Box 999, 99352 Richland, Washington, USA. [2]PsyberAnalytix LLC, 99352 Richland, Washington, USA.

Received: 17 January 2014 Accepted: 25 August 2014
Published online: 25 September 2014

References

1. CE Landwehr, Formal models for computer security. ACM Comput. Surv. **13**(3), 247–278 (1981)
2. DE Denning, *Information Warfare and Security*. (ACM Press, New York, 1999)

3. FL Greitzer, SL Clements, TE Carroll, JD Fluckiger, Towards a research agenda for cyber friendly fire. Technical Report PNNL-18995, Pacific Northwest National Laboratory (2009)
4. DH Andrews, KT Jabbour, Mitigating cyber friendly fire: a sub-category of cyber mishaps. High Front. **7**(3), 5–8 (2011)
5. United States Air Force Chief Scientist (AF/ST), Report on Technology Horizons: A Vision for Air Force Science & Technology During 2010–2030 (2010). http://www.au.af.mil/au/awc/awcgate/af/tech_horizons_vol-1_may2010.pdf
6. FL Greitzer, TE Carroll, AD Roberts, Cyber friendly fire: Research challenges for security informatics, in *Proc. of the IEEE International Conference on Intelligence and Security Informatics (ISI 2013)* (IEEE, Piscataway, NJ, 2013), pp. 94–99
7. FL Greitzer, J Strozer, S Cohen, J Bergey, J Cowley, A Moore, D Mundie, Unintentional insider threat: contributing factors, observables, and mitigation strategies, in *Proc. of the 47th Hawai'i International Conference on System Sciences (HICCS-47)* (IEEE, Piscataway, NJ, 2014)
8. S Leka, A Griffiths, T Cox, *Work Organization & Stress: Systematic Problem Approaches for Employers, Managers and Trade Union Representatives*, Protecting Workers' Health Series No. 3. (World Health Organization, 2004). http://www.who.int/occupational_health/publications/pwh3rev.pdf, Accessed 7 Dec 2013
9. P Lehner, M-M Seyed-Solorforough, MF O'Connor, S Sak, T Mullin, Cognitive biases and time stress in team decision making. **27**(5), 698–703 (1997)
10. E Soetens, J Hueting, F Wauters, Traces of fatigure in an attention task. Bull. Psychonomic Soc. **30**(2), 97–100 (1992)
11. BK Houston, Noise, task difficulty, and Stroop color-word performance. J. Exp. Psychol. **82**(2), 403–404 (1969)
12. AF Stokes, K Kite, *Flight Stress: Stress, Fatigue, and Performance in Aviation*. (Gower Technical, Aldershot, 1994)
13. DR Davies, R Parasuraman, *The Psychology of Vigilance*. (Academic Press, London, 1982)
14. GRJ Hockey, Changes in operator efficiency as a function of environmental stress, fatigue and circadian rhythms, in *Handbook of Perception and Performance*, ed. by K Boff, L Kaufman, and JP Thomas, vol. 2 (Wiley, New York, 1986), pp. 1–44
15. PL Wachtel, Anxiety, attention, and coping with threat. J. Abnorm. Psychol. **73**(2), 137–143 (1968)
16. MR Endsley, Towards a theory of situation awareness in dynamic systems. Hum. Factors. **37**(1), 32–64 (1995)
17. MD Rodgers, RH Mogfor, B Strauch, Post hoc assessment of situation awareness in air traffic control incidents and major aircraft accidents, in *Situation Awareness Analysis and Measurement*, ed. by MR Endsley, DJ Garland (Lawrenece Erlbaum Associates, Mahway, 2000). pp. 73–112
18. R Dhamija, JD Tygar, M Hearst, Why phishing works, in *Proc. of the SIGCHI Conference on Human Factors in Computer Systems (CHI '06)* (ACM, New York, 2006), pp. 581–590
19. D Sharek, C Swofford, M Wogalter, Failure to recognize fake Internet popup warning messages, in *Proc. of the Human Factors and Ergonomics Society 52nd Anual Meeting* (Human Factors and Ergonomics Society, Santa Monica, CA, 2008), pp. 557–560
20. J-P Erkkilä, Why we fall for phishing, in *Proc. of the SIGCHI Conference on Human Factors in Computer Systems (CHI '11)* (ACM, New York, 2011)
21. SG Hart, CD Wickens, Workload assessment and predictions, in *MANPRINT: An Approach to Systems Integration*, ed. by HR Booher (Van Nostrand Reihold, New York, 1990), pp. 257–296
22. S Dekker, M Lützhöft, Correspondence, cognition and sensemaking: a radical empiricist view of situation awareness, in *A Cognitive Approach to Situation Awareness: Theory and Application*, ed. by Banbury S, Tremblay S (Ashgate, Burlington, 2004), pp. 22–41
23. KE Weick, *Sensemaking in Organizations*, Foundations for Organiztional Science. (SAGE Publications, Thousand Oaks, 1995)
24. B McGuiness, L Foy, A subjective measure of SA: the crew awareness rating scale (CARS), in *Proc. of the Human Performance, Situation Awareness and Automation Conference* (SA Technologies, Marietta, GA, 2000)
25. N Sarter, D Woods, Situation awareness: a critical ill-defined phenomenon. Int. J. Aviat. Psychol. **1**(1), 45–57 (1991)
26. SWJ Kozlowski, Training and developing adaptive teams: Theory, principles, and research, in *Making Decisions Under Stress: Implications for Individual and Team Training*, ed. by JA Cannon-Bowers, E Salas (America Psychological Association, Washington, DC, 1988)
27. D Serfaty, J MacMillan, EE Entin, EB Entin, The decision-making expertise of battle commanders, in *Naturalistic Decision-Making*, ed. by CE Zsambok, G Klein (Lawrence Erlbaum, New York, 1997)
28. GP Tadda, JS Salerno, Overview of cyber situation awareness, in *Cyber Situational Awareness: Issues and Research*, ed. by S Jajodia, P Liu, V Swarup, and C Wang (Springer, New York, 2010), pp. 15–35
29. MR Endsley, WM Jones, Situation awareness, information dominance, & information warfare. Technical Report AL/CF-TR-1997-0156, US Air Force Armstrong Laboratory (1997). http://www.dtic.mil/dtic/tr/fulltext/u2/a347166.pdf
30. MR Endsley, WM Jones, A model of inter- and intrateam situation awareness: implications for design, training, and measurement, in *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*, ed. by M McNeese, E Salas, and M Endsley (Human Factors and Ergonomics Society, Santa Monica, 2001)
31. FL Greitzer, DH Andrews, Training strategies to mitigate expectancy-induced response bias in combat identification: a research agenda, in *Human Factors Issues in Combat Identification*, ed. by DH Andrews, RP Herz, and MB Wolf (Ashgate, Farnham, UK, 2010)
32. J Kavanagh, Stress and performance: a review of the literature and its applicability to the military. Technical Report TR-192, RAND (2005)
33. H Kohn, Effects of variations of intensity of experimentally induced stress situations upon certain aspects of perception and performance. J. Genet. Psychol. **85**, 289–304 (1954)
34. JA Easterbrook, The effect of emotion on cue utilization and the organization of behavior. Psychol. Rev. **66**, 183–201 (1959)
35. MA Staal, Stress, cognition, and human performance: a literature review and conceptual framework. Technical Report NASA/TM-2004-212824, National Aeronautics and Space Administration (2004)
36. IL Janis, L Mann, *Decision Making*. (The Free Press, New York, 1977)
37. IA Friedman, L Mann, Coping patterns in adolescent decision-making: an Israeli-Australian comparison. J. Adolesc. **16**, 187–199 (1993)
38. DR Davies, GS Tune, *Human Vigilance Performance*. (Staples Press, London, 1970)
39. J Leavitt, Cognitive demands of skating and stick handling in ice hockey. Can. J. Appl. Sport. Sci. **4**, 46–55 (1979)
40. MD Smith, CJ Chamberlin, Effect of adding cognitively demanding tasks on soccer skill performance. Percept. Mot. Skill. **75**, 955–961 (1992)
41. JP Van Overschelde, AF Healy, Learning of nondomain facts in high- and low-knowledge domains. J. Exp. Psychol. Learn. Mem. Cognit. **27**, 1160–1171 (2001)
42. M Allnutt, Human factors: Basic principles, in *Pilot Error*, ed. by R Hurst, LR Hurst (Aronson, New York, 1982), pp. 1–22
43. RP Barthol, ND Ku, Regression under stress to first learned behavior. J. Abnorm. Soc. Psychol. **59**(1), 134–136 (1959)
44. RB Zajonc, Social facilitation. Science. **149**, 269–274 (1965)
45. S Kuhlmann, M Piel, OT Wolf, Impaired memory retrieval after psychosocial stress in healthy young men. J. Neurosci. **25**(11), 2977–2982 (2005)
46. S Kuhlmann, OT Wolf, Arousal and cortisol interact in modulating memory consolidation in healthy yound men. Behav. Neurosci. **120**(1), 217–223 (2006)
47. JE Driskell, E Salas, J Jonston, Does stress lead to a loss of team perspective? Group Dynam.: Theor. Res. Pract. **3**, 291–302 (1999)
48. DA Norman, *The Psychology of Everyday Things*. (Basic Books, New York, 1988)
49. J Hollan, E Hutchins, D Kirsh, Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans. Comput. Hum. Interact. **7**(2), 174–196 (2000)
50. RJ Heuer Jr, Analysis of competing hypotheses, in *Psychology of Intelligence Analysis* (Center for the Study of Intelligence, Central Intelligence Agency, Washington, D.C., 1999)
51. MD Jones, *The Thinker's Toolkit: Fourteen Powerful Techniques for Problem Solving*. (Three Rivers Press, New York, 1998)
52. FL Greitzer, TE Carroll, AD Roberts, Cyber friendly fire. Technical Report PNNL-20821, Pacific Northwest National Laboratory (2011)

doi:10.1186/s13388-014-0013-5
Cite this article as: Carroll *et al.*: **Security informatics research challenges for mitigating cyber friendly fire.** *Security Informatics* 2014 **3**:13.