Original Article

A multi-phase network situational awareness cognitive task analysis

Robert F. Erbacher[a,*]
Deborah A. Frincke[b]
Pak Chung Wong[b]
Sarah Moody[a] and
Glenn Fink[b]

[a]Department of Computer Science, Utah State University, UMC 4205 Logan, Utah 84322, USA.
E-mails: Robert.Erbacher@usu.edu; s.j.m@aggiemail.usu.edu
[b]Pacific Northwest National Laboratory, PO Box 999, Richland, WA, USA.
E-mails: Deborah.Frincke@pnl.gov; Pak.Wong@pnl.gov; Glenn.Fink@pnl.gov

*Corresponding author.

Received: 1 March 2010
Revised: 13 June 2010
Accepted: 16 June 2010

Abstract The goal of our project is to create a set of next-generation cyber situational-awareness capabilities with applications to other domains in the long term. The objective is to improve the decision-making process to enable decision makers to choose better actions. To this end, we put extensive effort into making certain that we had feedback from network analysts and managers and understand what their genuine needs are. This article discusses the cognitive task-analysis methodology that we followed to acquire feedback from the analysts. This article also provides the details we acquired from the analysts on their processes, goals, concerns, the data and metadata that they analyze. Finally, we describe the generation of a novel task-flow diagram representing the activities of the target user base.
Information Visualization (2010) **9**, 204–219. doi:10.1057/ivs.2010.5

Keywords: cognitive task analysis; cyber security visualization; security analyst feedback; task-flow diagram

Introduction

Large-scale networks continue to exacerbate the problem of network management and identifying and isolating attacks, especially sophisticated attacks. Networks designed for large-scale connections, such as national and international databases, have additional issues associated with managing the network connectivity. The many connections made by legitimate users attempting to connect to local resources obfuscates attacker activity. While network managers must identify and eliminate malicious activity, they must simultaneously make certain that valid users have access. Given the enormous numbers of events that must be analyzed and classified, many solutions simply cannot scale to resolve all events deemed malicious.

Our goal was to develop novel techniques to aid in analyzing and interpreting this massive amount of data. The chaotic nature of network traffic data makes it quite difficult to accurately differentiate normal from malicious traffic. The network managers' goal is to prioritize the events based on their likelihood of maliciousness and potential ramifications of the event should it prove to be malicious.

To handle the ever-increasing numbers of attacks, network analysts and managers have processes and analysis stratagems for dealing with typical cyber attacks. Their first level of analysis is at a highly abstract, situational-awareness level. Only when an attack is identified at this level do they drill down into more detailed visualization techniques for actual analysis.

The goal of situational-awareness visualization for cyber analysis is to provide perceptually based displays that allow decision makers to rapidly understand the readiness of all available cyber resources. Readiness in this context is the ability of cyber resources to perform day-to-day tasks and deploy cyber operations and effects should they be designated to such activities. Existing situational-awareness environments, such as VisAlert,[1] lack representation of vulnerability assessment and impact analysis. These two components are critical for decision makers to accurately comprehend the state of the cyber resources.

We have created a set of next-generation, cyber-situational-awareness capabilities that, in the long term, will apply broadly to other domains. Situational awareness is the creation of abstract higher-level representations of the underlying raw data. It focuses on immediate comprehension rather than detailed analysis. Situational awareness is '… knowing what is going on so you can figure out what to do'.[2]

For situational awareness, we used Endsley's model.[3] This model intrinsically consists of three levels: perception, comprehension and projection. Perception is providing a representation of the current state of a situation. Comprehension relates to a higher-level understanding of all available data. Comprehension requires a far greater level of correlation and data integration than is incorporated into the perception level. Finally, the projection level looks at projecting the event into the future to determine its impact and progression.
The goal with situational awareness is to rapidly answer:

- What is happening?
- Why is it happening?
- What will happen next?
- What can I do about it?

Prior research has focused primarily on level 1 of Endsley's situational awareness model, namely providing for perception of events. We focused this work on both level 1 and level 2 of Endsley's model, adding extensive capabilities for comprehension through impact and vulnerability assessment. In additional, we developed examples of how level 3, projection, can be supported in new visualization designs. The goal is to improve the decision-making process such that better actions are taken. To satisfy the expectations of Endsley's situational awareness model and to verify that we were developing genuinely needed and useful capabilities, we performed the extensive cognitive task analysis (CTA)[4] described in this article. The CTA helped us understand the processes and needs of analysts and decision makers before designing the visualization techniques.

This article discusses the steps of our cognitive task analysis and shows the multiple phases at which we involved the experts in the process. Second, this article shows how we involved the experts, such as through developing scenarios to garner more useful feedback from the analysts. Third, we present the actual feedback from the analysts that will be instrumental in aiding researchers to develop more useful and effective techniques. Finally, we developed a new task-flow diagram outlining the idealized process followed by analysts and decision makers. This task-flow process identifies new capabilities needed for some of its phases to be followed more completely and effectively.

Previous Work

Existing situational-awareness capabilities have proven effective for representing and presenting situational data. However, these techniques are limited in their representation of only the status of the environment. They do not incorporate all of the characteristics needed by decision makers. We developed new visualization techniques that incorporate vulnerability and impact analysis on top of the base situational data to enable the interaction needed for exploration, analysis and effective action.

For the effective representation of situational awareness, we worked with experts from Pacific Northwest National Laboratory (PNNL) and the Air Force Research Laboratory (AFRL) to identify the characteristics that cyber command and control decision makers must be most readily aware of to effectively manage cyber resources. The environment must be adaptable to handle changing requirements by decision makers. Such changes will occur with the release of new attacks, identification of new vulnerabilities, deployment of new operations or cyber effects, changes in levels of hostility or threats by both external and internal agents and so on. One of the key results of working with these experts was to develop the task-flow diagram exemplified in Figure 1. This diagram became the roadmap from which we developed the visualization technique designs. 9+ This visual representation must provide an overview with details available on demand when decision makers need to view that information to identify an appropriate response. Finally, the decision maker must have the ability to identify when a situation has been resolved so that the situation can be removed from the display, and additional situations can be dealt with. This is critical for allowing decision makers to handle large environments effectively. Our research process is extended from our previous work in Endsley.[5]

This research examines issues related to the process described in this previous research, developing techniques to aid in the cyber analysis process. The process described in this previous work covers our research process from data acquisition through legal presentation. We must continuously consider the impact that the potential of legal action resulting from the analysis has on our techniques. Of particular importance is the need to show the analysis steps to a potential jury should legal remedies be sought. Existing cognitive task analyses have been done in related areas that provided a starting point. Although these studies have not been published, we have access to them through previous research efforts with the Air Force. Existing cognitive task analysis includes D'Amico *et al*[6] and D'Amico and Whitley,[7] which '… interviewed and observed IA analysts responsible for various aspects of cyber defense in six organizations within the U.S.

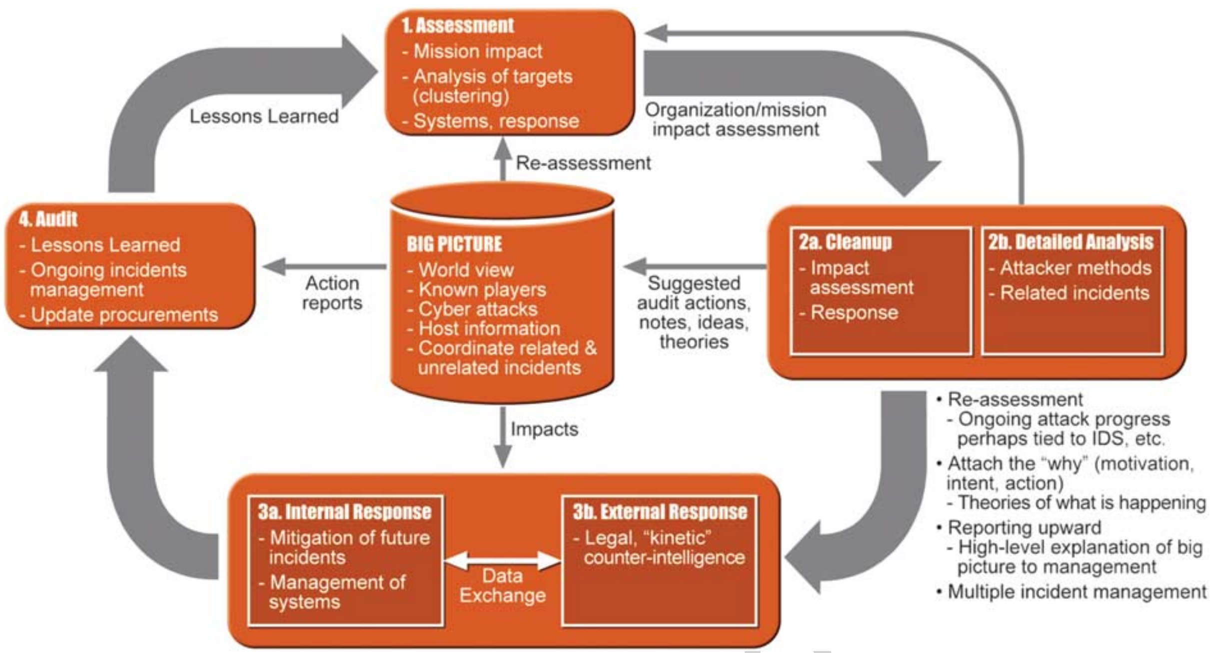

Figure 1: This task-flow diagram derives from our cognitive task analysis with Pacific Northwest National Laboratory.

Department of Defense (DOD) and one commercial managed service security provider (MSSP)'. This is essentially the forensic analysis of network intrusions.

Foresti and Agutter[8] conducted a CTA:

> To assess the work related demands of the Network Security Expert, an in depth description of the physical and mental workplace or a task analysis was undertaken. In order to construct or alter a system that supports the individuals decision making process, a detailed record of what the individual does or should do in the execution of their task is essential.

This work provided insight into data mining and decision support of network security experts.

Much of the analysis process is inhibited by the volume of data and the need to filter or cluster data. The extensive amount of data examination required makes digital data analysis a very slow process. Therefore, visualization techniques must be designed to work with typical forensic analysis capabilities. Through visualization and graphical user interfaces, the analyst's usual textual commands are still available but are made more accessible. We have already performed extensive but isolated aspects of visual-based analysis on which to build our new work.[5,8–10]

Even with the advantages of visualization in the analysis of large volumes of data, handling this scale of data is still a challenging problem for the field and remains an open problem. Techniques such as context and focus techniques,[11] clustering,[12] selective filtering and multiple views,[13] and aggregation[14] would all be applicable. For instance, the analyst could choose to display select event types needed for a particular analysis or aggregate events into event streams. The specific techniques deemed applicable will be determined by the exact application of a developed environment.

Methodology

The goal of our CTA study was to identify network managers' and analysts' specific needs in order to create next-generation visualization techniques. Thus, the CTA consisted of multiple steps in acquiring feedback from these experts. There were nine phases to our visualization design process; seven of these were solely focused on elements of the CTA:

1. CTA Phase 1: Initial brainstorming session
2. CTA Phase 2: Individual interviews
3. CTA Phase 3: Review of previous work
4. CTA Phase 4: Scenario creation
5. CTA Phase 5: Detailed brainstorming session and task-flow diagram formulation
6. Initial visualization design phase
7. CTA Phase 6: Visualization design review
8. Visualization re-design phase
9. CTA Phase 7: Program manager review

These phases are described in detail in the following section.

Results

For each phase identified in the above methodology, we discuss the characteristics of the participants and the major results acquired from the participants. Where possible, we grouped the participant thoughts based on our own interpretation of the goals and needs of the participant results.

CTA phase 1: Initial brainstorming session

The initial brainstorming session included network analysts, network managers, security researchers and visualization researchers at PNNL. This session was guided by the fundamental principles of the project. The principal goal of the project was to improve the situational awareness process, not by visualizing individual events, of which there are too many to be effective, but rather by identifying more effective, abstract concepts to be visualized. This brainstorming session identified key issues that we, and any visualization designer, would need to consider early in the visualization design process for this domain. The participants identified five potential primary goals for a situational awareness environment:

1. Impact identification. There are two main areas of impact: mission impact and system impact. Impact, especially mission impact, can be thought of as readiness in military terms.
2. Identification of the amount of damage.
3. Recovery.
4. Prevention and future projection. 'What-if' analysis in particular will be valuable in determining potential vulnerabilities and impacts from potential attacks.
5. Identification and characterization of attacks and the attackers.

Our requirements were to identify:

1. The malicious actors.
2. The legitimate users.
3. Abnormalities and subsequently the compromised systems.
4. The intended target of the attack, through trace back.
5. Main resources.

It is often hard to distinguish the good from bad usage of the system. Part of the reason for this difficulty is that our value judgments need to be as adaptive and dynamic as our adversaries' are.

In additional, extensive discussion centered on several other issues of importance to network managers:

Representation of subnets: Does the current plan assess vulnerability on subnets? Can you divide systems into smaller chunks to handle them easier? In particular, can the visualization handle large networks with multiple subnets? A system/subnet is as vulnerable as the most vulnerable piece found within it. Is the combination of systems done as an addition, multiplication, minimum/maximum, or floor/ceiling? To answer these questions, analysts need to know the assumptions being made by the software. It is more appropriate to know the assumptions being made by the software than to try to dictate the assumptions *a priori*.

Risk assessment: This should include identification of where the domain is not protected enough and what can be reached via a single breach. The visualization must also identify the impacts of the breaches – what data can be compromised and how network operation will be affected.

Support of network managers: Network managers supervise the analysts who respond to network attacks. What is interesting to look at will depend upon who the viewer is and the viewer's perspective within the organization. The network manager needs a higher-level (and different) view than a network analyst or system administrator. What the user is going to find most helpful will depend upon the particular user and his or her particular job and mission goal. To determine what type of capability will be more useful (for example, an alerting system, an analysis system, or a reporting and auditing system), the visualization must understand the various perspectives of the different users.

Support for templates: The system should support the ability to set up templates to handle different situations, such as

- life-threatening and other critical situations,
- varying levels of criticality,
- non-critical situations,
- logical versus physical representations,
- method for showing relationships and relativity, and
- what-if scenarios.

Templates will also aid in identifying what is normal. Future research must provide a means to validate templates so that users may be confident that they will always work. Human minds do not store information as templates. Thus, researchers need to develop concepts for what is appropriate for templates and how templates can most effectively be used and interpreted correctly. For instance, humans can easily recognize a chair or a chessboard. How can templates be developed such that they are as easily recognized?

Support communication: A complete command and control environment must support communication in a wide array of forms. A mechanism must be provided for analysts to communicate with each other. This is an issue of passing discoveries and expertise among the analysts and to the network manager for resolution identification.

For instance, more experienced network analysts will have more intuition and expertise in the analysis of network events. A less experienced analyst may identify events of interest but may not understand what they mean. In this scenario, they must be able to easily pass the results and details of their analysis to a more experienced analyst for examination without forcing the more experienced analyst to start from scratch.

Similarly, the communication capability can be used to monitor the resolution of an attack and verify that the resolution plan is followed. This will also help capture what was learned from the attack and provide after-action review, auditing and tracking.

Support for detection of sophisticated attacks. The environment should provide direct detection of longer term and more complex and sophisticated attacks. The assumption is that there is another level of detection being done at the triage level that likely would not detect low and slow sophisticated attacks. The visualization environment must look at dynamics over time to assess uncertainty.

Support interaction. Ultimately, interaction techniques are needed to focus the environment more on visual analytics techniques. The visual analytics capabilities are needed to analyze and interpret the data. This follows with the forensic goal of the environment as well.

At the end of the brainstorming session, several unresolved questions remained, such as the time frame needed to be considered for this level of analysis. Most of these questions were subsequently answered through existing cognitive task analysis, especially Anita D'Amico's from Secure Decisions[6,7] and Stefano Foresti's from the University of Utah.[8] Additional feedback was provided by the project manager at AFRL about the desired direction of the project. More specifically, it was identified that the triage period should be on the order of minutes. The forensic process begins after this and covers a period of hours to weeks. The goal of this environment was to support the post-triage period, handling the results of the attack in a more long-term forensic point of view. Additional questions that were unresolved include:

- Can analysts and researchers use attack trees? Can analysts visually identify the phases of an attack so that they can recognize what is happening before the last phase of the attack is completed?
- Are the visualization techniques incorporating baselining and identification of normal behaviors and so on? Are the visualization techniques looking at ports, running software, operating systems, networking and so on?
- What is planned for identifying the confidence level of the system information, that is, identifying how sure the analysts are of the validity of the data and their accuracy?

CTA phase 2: Individual interviews

The individual interviews with network managers and analysts focused on six subjects. We grouped most of the comments made by these subjects into the following major categories to ease interpretation and organization of the participant results:

- Process
- Management
- Communication
- Challenges
- Goals
- Data Sources and Querying
- Visualization Technique Requirements
- Visualization Organization
- Reference Tools

The following sections and tables are organized around the individual participants and their comments related to each of these categories. Tables 1, 2, 3, 4 contain the results from participants 1–4.

Participant 5

Management: Should economics be incorporated into the visualization? Can it help in the decision-making process? This becomes particularly relevant when measuring the cost of damages and determining whether to bring in law enforcement. Related information must similarly be considered for incorporation into the representation for decision making:

- What should be cut off or left out?
- What is the criticality of the identified item?
- Is it worth it to recover or just wipe the compromised system?
- Can all compromised points just be disconnected or are the systems important enough to keep up and running despite being compromised?
- Do the decision makers take the local domain completely off the internet or just close specific ports and access points?
- What is the effect of all these actions on cost and criticality of the situation?
- If it is a counter-intelligence (CI) situation, maybe the decision makers just doctor the data and let the attackers have them.

The response by decision makers can often be dictated by external situations or stimuli. This might include the existence of a foreign nexus, the case being handed over to CI personnel, malicious or disciplinary motivations and so on.

Challenges: One challenge is related to auditing and tracking. This becomes particularly difficult because so much communication among analysts and decision makers is done through email.

There is a need for trusted visualizations. This would require incorporating trust into the visualization and possibly forming a 'network of trust'. Consider, for example, domain name systems (DNS). Suppose information is trusted from some specific domains but not

Table 1: Participant 1's individual interview feedback for CTA Phase 2

Participant 1			
Communication	*Data sources and querying*	*Visualization techniquerequirements*	*Reference tools*
Analysts must be able to identify who is communicating with whom. This could be actual users, systems, applications and so on	Good log data are needed from as many sources as possible. A monitoring and logging tool that can query all logs in parallel would be good; consider web proxies	A tool is needed to show the network. For any visualization, one main goal is to get consolidated data (from firewalls, proxies, internet histories and so on) and then be able to filter out noise. This allows the analyst to see trends in the data. A sandbox browser is also a possibility, with sandboxes for each type of scenario, such as viruses	Consider the following tools for comparison: • Mojo Pack, U3 – thumb drive virtual machine • IPCAP – visual packet analysis environment • IN-SPIRE™ – visualization environment

Table 2: Participant 2's individual interview feedback for CTA Phase 2

Participant 2			
Process	*Management*	*Communication*	*Data sources and querying*
The process this analyst follows is the OODA loop developed by John Boyd:[15,16] • Observe • Orient • Decide • Act	In coordination with the Cooperative Protection Program, this analyst assists in the collection of data. The data are forwarded to the Operations and Analysis Center (OAC) or the Computer Incident Advisory Capability (CIAC) for analysis and correlation of context. This information is then sent on to a decision maker who determines what should be done and proceeds with the actions or sends it on to a responder who can take the appropriate actions	Use electronic Post – it® notes to share information about the data and results with other analysts and decision makers	It might be helpful to the visualization environment to provide the ability to filter data. This would either allow for the elimination of uninteresting data or limit the display to interesting data as well as related data

from any others. However, one of the trusted domains trusts other domains that is not directly trusted – how is information represented from this third domain? If it passes through the trusted domain to our domain, is that as good as it coming directly from the trusted domain – is it trusted or not? What if the information came directly from the third domain – is the information trusted? Do our systems automatically trust everyone that someone in our domain trusts? How can this be incorporated into a visualization where the information could be helpful?

Visualization Technique Requirements: Nine items identified by this participant need to be considered in the development of visualization techniques:

- A timeline. The timeline becomes critical for identifying the ordering of events and actions that have taken place and projected actions on what will take place. The timeline needs to be filled in as more data are acquired.
- Attacker entry points. It is important to identify the attacker's entry point and the likelihood of that entry point being used.
- Identification of compromised and infected systems. Who has been infected and what resources have been touched? Can the possible permutations of the attacker's actions be tracked? This goes back to having an accurate timeline to help identify when things could have happened and how far the attacker could have gotten. Because of the fuzziness of time and the fact that the attacker could have modified the logs, mirrors should be used to get a more accurate time frame of reference.
- Human presence and impact. There is always a human presence somewhere in the picture. For instance, someone designed the code that is performing the attack, someone designated the targets or at least the protocol for identifying targets and so on. Can this human presence be tracked? If the attack was introduced through a computer, where was the employee who uses that computer at the time of the attack? This information could provide quite a bit of detail that could be useful in handling the situation. Can this information be incorporated into a visual network model?
- Identification of what is on or off, up or down.
- Identification of system and network vulnerabilities, both physical and virtual. For example, the virtual space may have multiple redundant connections

Table 3: Participant 3's individual interview feedback for CTA Phase 2

			Participant 3		
Process	*Challenges*	*Goals*	*Data sources and querying*	*Visualization technique requirements*	*Visualization organization*
This analyst's rough pattern for incident response includes four phases as follows: 1. Identify 2. Eliminate 3. Clean 4. Restore	Some of the challenges faced are live or temporal data, the fact that attacks can happen over a long period of time, and the fact that intrusion detection is behavioral	To aid analysis, it must be determined if the attack uses common protocols or unusual or uncommon ones	The visualization should have a weight based upon accuracy of information: How much do analysts trust this information? Include automatic double-checking, ie, validation, of the data. Consider incorporating the following types of data into the visualization: impact analysis, protocols, authentication used, attack and attack code metadata, and source of attack or related resources. This might include URLs, filenames, class libraries, etc	Some things that would be good to incorporate into visualization for cyber situational awareness include: • The ability to design and submit queries • The ability to handle the huge data load • The capability to identify the victim machines quickly • A design to get the forensics on the hard drives started early • The capability to isolate the systems involved if possible • The ability to track the state of the various pieces of evidence • Incorporation of network location correlations • Incorporation of automated search facilities to search textual logs • Identification of who is talking to whom • Incorporation of behavioral norms. For example, is this computer supposed to be talking with that one? Answer: It should not be, usually does not need to, or it never has in the past • Incorporation of a model on how attackers got into the network in the first place, such as a Secure Shell (SSH), Secure Socket Layer, or SSH tunnel. Identify their point of entry and why it did or did not work	Have a top-down approach and roll-up summaries

Table 4: Participant 4's individual interview feedback for CTA Phase 2

Participant 4				
Management	*Challenges*	*Data sources and querying*	*Visualization technique requirements*	*Visualization organization*
Some of the major problems with incident response teams include: • People are not available • Analysts and managers have limited understanding of the software • Each attack is unique to handle because of the incident, the systems involved, work being done, tools needed, etc • Many reports are useless; they are often too technical. The goal is to make technical details understandable to management	A current challenge is the modeling of malware, especially those that are introduced to a system through direct contact. These include things picked up from home or while on travel, from email attachments, and from compromised USB flash drives	All of the following are monitoring points and sources of data that need to be considered: switches, routers, intrusion detection systems, uninterruptible power supplies, network monitors, network software, etc It would be beneficial to be able to model the environment down to the most detail possible. Analysts want to know the location of every router, switch and connection, as well as the specifics of all the hardware in the network, specifics of what software is running where, and so on. This is a vast amount of detail to model	Visualizations are needed that: 1. Verify if the attack is 'for real' 2. Identify whether the threat is an insider or an outsider source 3. Identify how to limit damage and impact (a) If it is an insider threat, then there is no benefit to cutting ourselves off from the outside (b) Set monitors that monitor outgoing information • Identifies what data are going where • Prevents compromise of classified data (c) Identify the attacker's intent • Are they after one machine or a mass of them? • What is the meaning behind the attacks? • What set it up? (Looking at code). For instance, what allowed it to happen? (d) Compared to known malware; is it the same or similar? (e) Identify malware and how it was introduced This allows analysts to: • Identify holes • Model malware and introduce it to the network model to see what happens and find solutions • Manipulate the model to improve security, justify budget requests, provide malware impact analysis, etc This could then be something to take in and show management. For example: this is what happened, but this is how bad it could have been if it had not caught it for another 2 days; this is how bad it was, but it could have been better if we had had software *X* running; this is why we did what we did and why it worked and why these other actions would not have worked	The visualization techniques must not only model the network but also model the security of the network. With just a model of the network, details and information about how something happened are lost

from a critical node to another node, but if they are mapped onto the physical network, they all go through the same physical box. This vulnerability can be detected by looking at the physical overlay and including the situational context. What are the single points of weakness? What are the communication points?
- Automatic updating for a systems upgrade, or new versions of the tool.
- A tutorial on how to get started; not just the user's manual, but a quick-start guide that can get a group up and running quickly.
- Provision of training on how to most effectively use the tool. Perhaps have a certification process so people can become certified to use the tool rather than just figuring it out on their own with a huge user's manual.

Visualization Organization: Incorporate the timeline into an overlay. Use overlays instead of filters; overlays let you see multiple things simultaneously more easily. For example, in a battle situation, a commander wants to see where his ground troops are and then add the information about artillery. This can be shown as a separate layer in a three-dimensional (3D) visualization. The commander will want to know the paths for the artillery shells as well as where they are going to hit so they do not hit the ground units. Additional overlays can add in air units and their flight patterns, making sure that the flight paths do not cross an artillery path when the artillery shell will be there. This visualization necessarily requires a 3D viewpoint factoring in time. Users want to be able to pull back some layers and look down through the rest for the entire picture. Users also need the ability to manipulate the entire viewpoint. In terms of network security, typical overlays might include:

- The trust model. As discussed in the section 'Participant 5', this relates to how reliable the data we are receiving are. Can the data be validated? Where are the data coming from and how trusted is that source. Techniques such as force-directed graphs[17] where trust is the primary spring force or, uncertainty visualizations[18] are examples of traditional visualization techniques applicable to the representation of trust.
- The system importance and criticality.
- Sensitivity model.
- Defensive capabilities, which should include granularity, identifying at what levels defensive capabilities exist.
- Use of different overlays for the physical and the virtual world and network. Without the physical overlay, the analyst is reduced to following the physical connections manually. It is time-consuming to follow wires to figure out where they go.

Reference Tools: A visualization tool currently in use is Big Brother (bb2).

Participant 6
Process: Some of the most important actions during an incident are collaborating, collecting and disseminating lessons learned as well as identifying the attack.
Management: A complete environment needs to provide a list of what to do, like a reaction plan. Similarly, a log of actions already taken, completed or in progress is needed. This log of actions needs additional capabilities to be useful and effective:

- Make this fluid and flexible to account for the uniqueness of each incident.
- Identify what parts can be done in parallel.
- Contact vendors and other parties involved.
- Determine if legal proceedings will take place, spawning off the workflow to legal authorities and decision makers for them to determine this. Currently, there are too many challenges to get people to prosecute. It is not advantageous for companies to prosecute because they would have to admit they were hacked, which is not good for business. Companies may need to be provided with incentives to pursue prosecution of electronic crimes. This might include tax breaks or penalties for not sharing data and so on. Companies lose more by prosecuting than they do by settling independently with the hacker because they can protect their credibility by not prosecuting.
- Determine if it is feasible to allow automatic contacting of those involved or who should be informed of the incident (those on a need-to-know basis).
- Dictate a clear-cut chain of command. This must be identified in correlation with the type of event.
- Support collaboration meetings to make sure everyone is on the same page. Such meetings may occur daily during an incident.
- Identify remediation steps. At the management level, this would include identifying the impact, providing public releases, answering questions and having after-action reviews with lessons learned. Lessons go to all incident response teams so that everyone can learn from the situation rather than just the team who dealt with the incident. It also helps identify ways to improve efficiency of the entire process.

Communication: It would be useful to incorporate some sort of communication medium within the visualization tool, like a board, to share data. In some cases, analysts have used a whiteboard, markers and Post-it notes. Incorporate an electronic board, such as a sophisticated discussion board. Such a paradigm could help with briefings such as daily updates, to do lists and so on.

An additional communication board is needed to relate information to management. This could include the dynamic loading of status and details to a website so management could be kept up to date via their mobile devices. Such information can easily be protected in a 'management only' section to avoid many of the technical details. Similarly, this presentation could provide summarizations of actions taken and a single point of contact for an event.

Goals: The ability to judge the full impact of an event is needed. This might include the following components:

- How many nodes or systems were hit or compromised?
- What type of data were compromised? Were they Non-Disclosure Agreement data?
- What data were exfiltrated? Were they proprietary?
- What kind of vulnerability was used (call home, back door–reverse tunneling)? Was it widely known or in news stories? If this is the case, it is bad because it indicates susceptibility to common well-known vulnerabilities.

Data Sources and Querying: It is important to know what triggered the incident, whether it was specific or generic. Analysts must also know how relevant each piece of datum is. The use of weighted importance will be helpful in this, allowing users to weight information and include a description of their reasoning behind the weighting. This weighting can also be tagged with information as to who increased/decreased the importance and why, and possibly what and where.

To obtain an accurate picture of situational awareness, analysts also need to know what is going on outside, that is, external information. For instance:

- Is our domain the only one being attacked or are others being attacked?
- Is this attack common or rare?
- Is it a new attack or has it occurred before?

This information helps to provide what is actually needed, situational awareness of the entire network.

Specific data and metadata that need to be incorporated include:

- People
  - Who owns the data?
  - Was it personally identifiable information?
  - Who takes care of what?
- What is the time frame of the event?
- Source Internet protocol (IP)
- Destination IP
- Packet Capture (PCAP)
- Ports
- Date information. More specifically, when was the information added?

Visualization Technique Requirements: A clear identification of the assumptions that the environment is working under is needed. It is not as important if the user can designate what assumptions should be made. The goal is to incorporate system figuring, that is, identify information about the system, such as what the environment knows and assumes about the system.

The visualization needs to incorporate current policy because security is defined by the policy. Knowing the next action according to policy will help in the decision-making process. This visualization of the organization's policy could be added as an additional filter or layer that could be shown or not shown, depending upon the analysts' preferences and selections.

Having some form of geolocation integrated into the visualization environment would be useful. This integration may allow the analyst to track the event back to its source. It will also allow the scale of the attack to be identified.

Finally, providing the ability to view a video or simulation of the events will be helpful in reviewing what occurred more rapidly. This ability can also help bring new analysts and managers up to speed on a scenario.

Visualization Organization: Three things that should be incorporated are (1) a representation of the generalized attack path, (2) a representation including all nodes and routers, and (3) a representation of a timeline of events.

CTA phase 3: Review of previous work

This phase of the research reviewed existing techniques that might be applicable to this domain. This phase focused on techniques that the analysts identified as being of interest. We examined previous visualization techniques that applied overlays and different communication board techniques.

CTA phase 4: Scenario creation

Six scenarios were created based on the background and requirements acquired to this point. These scenarios were put together to elicit feedback from network experts as to their workflow and needs with respect to potential visualization activities. These scenarios were chosen to be representative of a variety of different situations and conditions that the analysts would typically be required to deal with. The scenarios were then presented to visualization experts to elicit additional ideas and concepts for needed visual presentations. Below, we describe the six individual scenarios as well as our initial impression of the requirements of a visualization system to handle these scenarios. The section 'CTA Phase 5: Detailed Brainstorming Session and Task-flow Diagram Formulation' identifies the feedback acquired from the analysts about these scenarios.

Scenario 1: Large-scale disparate attack

A large-scale cyber attack is detected against military targets. The attack appears to be from a new terrorist group located in a third-world country. The group does not appear to be very sophisticated, but it is gaining sophistication with training and guidance from the People's Liberation Army. The group is using a wide range of different attacks, ranging from script kiddies to attacks that are more sophisticated. Thus far, there have not been any zero-day attacks, but the group is using just about every other attack.

Visualization requirements:

- What systems are potentially the most vulnerable?
- What systems and networks are being attacked?
- Who has been notified of the ongoing attacks?
- What is the status of the attacked systems?

What remediation for the attacks has been initiated: none, firewall updates, network disconnected, systems disconnected and so on.

Scenario 2: Distributed denial of service attack
A distributed denial of service attack is being applied to a critical sub-network, specifically attacking the payroll and human resources facilities of Fort Barksdale. The goal is to create dissention within the cyber capabilities of the United States and the US Air Force.

This is an extensive denial of service attack using a large number of systems compromised over a number of years, that is, sleeper systems. The attack is using multiple routes and attacking all publicly known network access points.

Visualization requirements:

- Exactly how much external bandwidth remains, especially to critical remote systems, for example, financial institutions?
- To what extent can payroll and human resource services function; for example, how quickly?
- Have the appropriate personnel of authority been notified?
- Have alternative mechanisms been arranged?
- Where is the remote activity coming from?

Where can the external activity reasonably be blocked at its earliest point to reduce its impact?

Scenario 3: Brute-force attack leading to a successful system compromise
A brute-force password attack has compromised an internal system. The system was then remotely connected to. Files were uploaded onto the compromised system, and the system began scanning other systems within the local subnet.

Visualization requirements:

- Identification of the numerous remote accesses indicative of the attack.
- What is the status of the compromised system: disconnected, shutdown, operating system reinstalled, online and so on?
- What was the system used for? What functionality was lost? And to what extent?
- To what extent was the system compromised: only the password, rootkit, application compromise, data exfiltration and so on?
- What secondary systems were scanned?
- What is the status of the secondary systems?
- What other system(s) had/have the same user account?
- What other system(s) has/have that user logged into recently?
- What is the status of the event: new, analyzed, resolved and so on?

Who has been notified/taken responsibility: network manager, system owner, security and so on?

Scenario 4: Kernel buffer overflow attack leading to a successful system compromise
A bug in the Windows firewall software allows an attacker to apply a buffer overflow attack to gain access to the kernel and thus generate a remote login. A remote attacker gains access to one such system.

Visualization requirements:

- Identification of the remote access indicative of the attack.
- What data were transmitted to initiate the attack?
- What other systems are vulnerable to this attack?
- What networks/subnets/systems have had defenses added for this attack?
- What is the status of the compromised system: disconnected, shutdown, operating system reinstalled, online and so on?
- To what extent was the system compromised: only the password, rootkit, application compromise, data exfiltration and so on?
- What secondary systems were scanned?
- What is the status of the secondary systems?
- What other systems had/have the same user account?
- What other system(s) has/have that user logged into recently?
- What is the status of the event: new, analyzed, resolved and so on?

Who has been notified/taken responsibility: network manager, system owner, security and so on?

Scenario 5: Zero-day worm attack
A worm applying a new vulnerability of Apache web servers is spreading rapidly over the internet. To maximize the impact, the attackers infected DOD systems first.

Visualization requirements:

- What systems are running an Apache server?
- What is the status of the compromised system: disconnected, shutdown, operating system reinstalled, online and so on?

- What is the status of the vulnerable systems: Apache disabled, system shut down, disconnected, new firewall rules and so on?
- To what extent was the system compromised: Apache compromised, rootkit installed and so on?
- What secondary systems were scanned?
- What is the status of the secondary systems?

Who has been notified/taken responsibility: network manager, system owner, security and so on?

Scenario 6: DNS cache poisoning
An internal subnet has been compromised through DNS cache poisoning. This has redirected traffic for internal email and an internal website to an unauthorized system.

The visualization should show the misdirected traffic as well as event statistics.

Visualization requirements:

- Where is internally designated email and web traffic going?
- How much data were compromised and from what sources?
- What other systems are vulnerable to this attack?
- What types of data were compromised?
- Who has been notified of the event: network manager, system owners and individual users?

Where were the data being redirected to: internal system, external system and so on?

CTA phase 5: Detailed brainstorming session and task-flow diagram formulation

The scenarios described in the section 'CTA Phase 4: Scenario creation' were presented to a variety of analysts. After their initial feedback on these scenarios was acquired, a discussion of their task-flow in such scenarios followed. This follow-on discussion led to the creation of the task-flow diagram shown in Figure 1. The analyst feedback follows. The full explanation of the task-flow diagram is provided in the section 'Task-flow diagram'. Given this was the most detail oriented of the task analysis phases, we document the background of the participants in Table 5.

Scenario 1: Large-scale disparate attack
This is a viable scenario, but a military origin is not required. The environment needs to show what is actually being attacked (is it just web servers and so on?). It would be helpful to have some kind of underlying artificial intelligence that can make generalizations and inferences about the attacks. The level of commonality among systems needs to be shown; it is assumed that the analysts know that an identifiable, large-scale attack is underway. For instance, are there consistencies between the systems being attacked that point to a common vulnerability? The visualizations should help identify a concerted attack from the typical background noise. Some specific ideas to incorporate into visualizations for this scenario might include:

- Initially, provide a target-oriented view (How are the targets related?)
- Next, the visualizations should simultaneously allow the analyst to:
  - Analyze the attacks if funding and time are available using a simulated environment. This analysis would require the following steps:
    - Quarantine and forensically analyze data.
    - Image the affected machines into a virtual environment.
  - Perform clean up and containment, including patching, applying countermeasure and so on. This task may require taking the systems off-line quickly, reformatting them and reinstalling the operating system.
- It would be worthwhile to investigate providing techniques to visualize progress – more specifically, visualizing the workflow of the analysts, such as their tasks and progress.
- Visualization techniques are needed to share analysis across sites and manage information sharing during the analysis.
- A complete environment needs to provide auditability. This is important in general but also becomes important for training and education.
- The impact of law enforcement involvement and the requirements it places on the visualization environment must be considered.
- Phased analysis is important and must be supported in the environment.

Scenario 2: Distributed denial of service attack
Can analysts filter out the distributed denial of service traffic so they can see what else is going on?

Analysts must be able to look at impacts from a non-computational, mission-oriented perspective. They must be able to consider the mission impact from any response they initiate.

The desired solution will be more than just a new visualization environment or technique for incident management; it will be a command console plus decision-support system. The visualization environment should be designed to plug into a larger command and control system.

The ease with which the analyst can inject data into the system is important. Ultimately, there will need to be automated processes feeding status information back to a monitoring database. Subsequently, it is valuable to

Table 5: Cognitive task analyses Phase 5 participants. All participants are associated with PNNL

Name	*Title*	*Role*
Participant 1	Scientist Secure Cyber Systems Group	National Security Directorate, Computational & Statistical Analytics Division. In addition, he is the Information, Infrastructure and Integrity Initiative (I4) focus area lead for projects relating to adaptive systems. His primary role is to create revolutionary cyber security technologies, especially in the area of adaptive systems. His primary expertise is visualization, and he has general expertise in human–computer interfaces
Participant 2	Manager Secure Cyber Systems	National Security Directorate, Computational & Statistical Analytics Division. He is the deployment manager for the Radiation Portal Monitoring Project for the wireless for airport vectors. He serves as program manager and network engineer in several projects
Participant 3	Scientist W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL)	Staff member in the organization of the Associate Director, Enabling Technology. He leads and participates as technical staff on scientific desktop, infrastructure, security and high-performance computing cluster teams, providing support for over 1000 Unix/Linux/Mac/Windows systems, security infrastructure, including four Nessus security scanners located in different security zones, and over 20 compute clusters providing over 1500 CPU cores for scientific data processing. His security team provides the capability to audit and assess networks using Nessus and is funded to perform internal PNNL assessments on a bi-yearly basis. He also performs/participates in assessments for external customers Capability Steward for Computer and Network Services (CaNS) in the Enabling Technologies Directorate of EMSL. He manages a US$1.5 million budget. He maintains CaNS capabilities at a state of readiness to assist EMSL staff and users with their information technology (IT) needs. He makes strategic investments to improve efficiency or add new capabilities that facilitate scientific discovery. He provides ongoing maintenance of CaNS systems as necessary to keep EMSL capabilities working properly. He manages the allocation of CaNS resources (people, equipment, parts and laboratory space). He makes available knowledgeable staff, supplies, and workplaces for EMSL staff and visitors
Participant 4	Scientist EMSL	Staff member in the organization of the Associate Director, Enabling Technology. He works extensively on classified and unclassified vulnerability remediation as well as internal and external audit planning and execution
Participant 5	Scientist Information Analytics Group Research Coordinator for National Visualization and Analytics Center	National Security Directorate, Computational & Statistical Analytics Division. He is a research scientist at PNNL, specializing in knowledge representation, visualization, collaborative tools and information integration. He has designed knowledge sharing and collaboration capabilities. He performs research in visualization design and implementation, analytic reasoning environments, knowledge representation and data-intensive computing for applications in information analysis and cyber security

provide annotations on attacker motivations and intent as well as theories as to what is or was occurring.

Often, administrators hear about attacks from other individuals. Can a system be provided that is always theorizing that one or more IP addresses are intruders? Can the system pretend that x.y.z.a is an attacker and identify what it is doing? Does this enable getting a jump on identifying actual attacks and attackers? This 'theorizing' system could have an enormous potential for false positives.

The environment must support ongoing assessment. Analysts cannot separate the first 2 hours of an attack from the 2-hour to 2-week time frame. Can analysts take notes and report a new incident, such as with a problem report?

Scenario 3: Brute-force attack leading to a successful system compromise

When developing the visualization environment, consider representing the level of commonality between compromised systems, that is, showing what characteristics they have in common. How can the visualizations compare systems?

Solutions are needed that can be extended to the physical domain. Consider, for instance, the impact of identifying a physical door breach in correlation with a cyber security breach. The analysis of the connectivity between physical and cyber security should include:

- Identify how physical security affects the organization's cyber security.
- The environment needs to visualize the cyber attack in a physical metaphor.
- Many of the cyber security problems have been solved in the physical world. Can analysts relate the two and learn from the lessons learned in the physical world?
- The visualization should create a picture of the attacker's theoretical workflow and allow for adjustment of expectations in real time. Consider hurricane tracking

for comparison. Can a similar type of forecasting for cyber attacks be created?

Scenario 4: Kernel buffer overflow attack leading to a successful system compromise
In this type of scenario, the visualization would need to correlate the activity with the fact that there is no previous valid login. Additional information could be acquired from BugTraq and so on. This additional data and metadata would then have to be visualized.

The environment should securely share the incident data with the rest of the analysis team without sharing the data more broadly. How auditable and legally defensible is this approach? Does making the audit record write-once affect people's willingness to record the event data? How about if the record were erasable, auto-signable, non-erasable or with tracking changes? Consider for example:

- A wiki without the wiki interface. Instead of the typical interface, enable a paradigm in which the analyst sends an email to the wiki and has the attached data automatically added.
- Provide the ability to drill down to memory dumps of the system, process tables and so on. Analysts could create a system image from a menu button.
- The wiki should track where analysts visit and provide feedback: 'most analysts who visited this site went < here > next'. This capability could greatly improve analyst efficiency.

Scenario 5: Zero-day worm attack
For this scenario, consider a four-phase process:

1. Perform mission impact assessment. Identify the attack's impact on systems, resources and so on.
2. Perform the following three responses in parallel:
   - Cleanup
   - Detailed analysis
   - Sharing information between cleanup and analysis teams
3. Mitigate the attack impact and potential future compromises, by:
   - Considering future mitigations
   - Pursuing legal remedies if actionable
   - Implementing a kinetic response.
4. Audit at the entire incident level. Though this level of auditing almost never gets done, it is sorely needed, and capabilities to simplify it would greatly help. Add the capability to allow administrators to mark things for later follow-up. Provide for report generation.

This process is very reactionary, but do analysts and decision makers care?

Scenario 6: DNS cache poisoning
No additional feedback was provided; this scenario includes characteristics from previous scenarios.

Initial visualization design phase

The results of the task-flow diagram identified the need for multiple levels of visualization techniques. Typically, visualization techniques for network analysts have focused solely on analysis. We developed techniques designed at the high level to provide the needed situational awareness and immediate assessment as well as detailed analysis and interpretation. This is exemplified by analysts needing to check the status of the network first thing in the morning for an immediate overall assessment. Should a problem be identified, then a more detailed analysis would be required to identify what was causing the problem and its source. Our visualization technique design was based on the need for these multiple levels of visualization analysis.

In additional, the analyst interviews identified specific features that would need to be incorporated into a complete visual analysis system. Many of these features would be considered future work and were out of scope of the current project, such as the need for a communication board.

CTA Phase 6: Visualization design review

The initial visualization techniques were reviewed by the analysts to identify the extent to which the designs appeared to meet their needs. Initially, six visualization designs were presented. Individual interviews were performed with six analysts. These were the same analysts used in previous steps of the CTA, and thus they were familiar with the project and the previous discussions. Their feedback, both positive and negative, associated with each of the visualization techniques was recorded. This process resulted in extensive documentation on the perceived benefits of the techniques. On the basis of the analyst feedback, two visualization techniques were chosen for further refinement. One of the chosen techniques focused on high-level immediate assessment (situational awareness) of the overall network status. The second chosen visualization technique was designed around a more detailed analysis paradigm. The goal is for the chosen visualization techniques to work together in real deployed scenarios.

While a discussion of the individual visualization techniques and the analysts' specific comments is beyond the scope of this article, the goal of showing the steps in a complete CTA and its impact on the development of visualization techniques for cyber situational awareness and analysis has been achieved.

Visualization re-design phase

The two chosen visualization designs were revamped to resolve the analyst-identified issues and to better document the behavior and appearance of the visualization techniques. We then implemented prototypes of the two redesigned visualization techniques. The visualization techniques were then tested with the Skaion data set. The Skaion data set was created with funding from the Defense Advanced Research Projects Agency (DARPA) and is released as Official Use Only. The Skaion data set is used by AFRL and other government agencies because it is more complete, more representative of real network activity, and does not contain any actual sensitive data. The Skaion data set does not contain any of the known limitations of other data sets, such as the Darpa/Massachusetts Institute of Technology data set.[19] There are four versions of this data set with approximately 1 GB, 3.5 GB, 3.2 GB and 50 GB of data in each set. For our testing, we used version 3 of the data set. This version included:

- Ground truth data
- Scenario description
- Network topology map
- Testbed description
- Testbed configuration description
- Network flow data
- Dragon data
- File transfer protocol (FTP) logs
- Snort data
- Stepping stones data
- Tcpdump data
- Unix logs
- Web logs
- Win logs
- Additional supplemental and support documentation.

CTA phase 7: Program manager review

The client program manager for the project reviewed the implemented prototypes using the data set. The program manager's impressions of the capabilities were positive. This review served as the final review of the project and developed techniques. For future follow-on work, we would like to perform quantitative user studies to fully assess the capabilities and effectiveness of the developed techniques compared to both other existing visualization techniques for network analyses and existing textual tools used for this purpose. For the latter, it is the tools that are currently in use by network analysts that are of particular interest.

Task-flow Diagram

The primary characteristic of interest from the task-flow is the need for multiple visualizations to support the different task-flow phases exemplified in the diagram (Figure 1). For instance, the first phase, labeled *Assessment* in the diagram, requires that the domain expert be provided with a summary display of the overall network status. This assessment would need to support multiple levels: mission or task, system, targets, capability and so on. This display must provide for a rapid understanding of the capabilities at a broad level. For this project, we targeted the representation of impact and vulnerability assessment of systems for the summary displays. This summary representation is what really differs from current and previous work in situational awareness visualizations that have focused on detailed representations indicative of what is identified as needed for the detailed analysis phase within the task-flow diagram.

The Detailed Analysis, Phase 2, and the Big Picture require a second visualization framework. This visualization capability requires a more detailed representation of the network. It is here that the network infrastructure *must* be represented in some form. This is not to say that the network infrastructure must be the only representation, but it must be one option available to the analysts. This need was made very clear by the analysts we interviewed. This detailed display must include typical situational awareness capabilities but will also include:

- Correlation. What do attacked systems have in common and what other systems have this commonality? The commonality could be the same service, the same user or the same administrator.
- Attack strategies. What methods of attack are being used such that they can be defended against?
- Impact assessment. What users, capabilities, missions, or systems are being affected and to what extent?
- System criticality. How important is the system? This criticality assessment could be related to the importance of a particular system, a capability or a user of the system.
- Vulnerability assessment. What systems are vulnerable to a particular attack? What systems in general are more vulnerable to attack?

Phase 3, the Response phase, and Phase 4, the Audit phase, indicate additional capabilities required by the infrastructure as a whole. It is here that we have identified the need for more of a command console in addition to the situational awareness capabilities. The idea of a command console would include more event and incident management capabilities, for instance:

- Who has been notified of the incident?
- Who has taken responsibility for the incident?
- What is the status of the system?
- What must be done to resolve the incident?
- What type of attack was it? Why was it successful?
- What is the status of the event, and what progress has been made in its resolution and prevention?

Although Phases 3 and 4 were out of scope for this project, they identify capabilities the network analysts and managers need to improve for managing events in the long term. For instance, currently, many actions that are identified as the result of an incident are not followed up on because of a lack of event management facilities. For military purposes, this amounts to a lack of follow-through of tasks identified from after-action reviews.

Conclusions

We followed a human-in-the-loop research process, creating a new task-flow model for network management that will greatly affect future research in this domain. The analyst interviews identified the basic requirements, critical parameters and characteristics needed for the next generation of cyber situational awareness visualizations. These analyst interviews resulted in a new view of the needed capabilities for network analysts and managers, resulting in a new task-flow diagram representing such analyst activities.

The full description of the cognitive task analysis provides two additional benefits. First, the process we followed for the cognitive task analysis, especially the inclusion of experts in the visualization design process from start to finish, can provide insight to other visualization designers. Traditionally, visualization design has incorporated very little if any cognitive task analysis. Thus, this insight into how domain experts can be fully integrated into the process and what information to seek from them will aid other visualization experts in more effectively developing their own processes.

Second, the specific details acquired from analysts and supplemental data, such as the six scenarios we developed for discussion, will greatly aid researchers in the cyber security domain. In particular, it will aid cyber security researchers in identifying what capabilities are needed at both macro and micro levels. It also provides insight as to what data analysts typically have available and the processes they typically follow.

Because most cyber security researchers will not have analysts available for similar types of analysis, this research will provide the background needed for them to perform needed and more useful research.

Acknowledgements

This research was funded in part by AFRL under project FA8750-07-C-0163. Many students played significant roles in the performance of the project, including Anupama Biswas, Anusha Davuluri, Chris Harris, Srinidhi Kakani, Stephen Miller, Steena Montiero, Sarah Moody, Rian Shelley and RB Whitaker. We also appreciate the time and effort of the participants from PNNL who provided input during the cognitive task analysis. We also appreciate the valuable assistance and comments from Sharon Eaton, Lorie Layne and Lee Ann Dudney, from Pacific Northwest National Laboratory, in developing this article.

References

1 Foresti, S., Agutter, J., Livnat, Y., Erbacher, R. and Moon, S. (2006) Visual correlation of network alerts. *Computer Graphics and Applications* 26(2): 48–59.

2 Adam, E.C. (1993) Fighter cockpits of the future. Proceedings of 12th IEEE/AIAA Digital Avionics Systems Conference (DASC), Fort Worth, TX: IEEE Press, pp. 318–323.

3 Endsley, M.R. (1995b) Toward a theory of situation awareness in dynamic systems. *Human Factors* 37(1): 32–64.

4 Crandall, B., Klein, G. and Hoffman, R. (2006) *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

5 Erbacher, R.F., Christensen, K. and Sundberg, A. (2006) Visual forensic techniques and processes. Proceedings of the 9th Annual NYS Cyber Security Conference Symposium on Information Assurance; June 2006, Albany, NY, pp. 72–80.

6 D'Amico, A., Tesone, D., Whitley, K., O'Brien, B. and Roth, E. (2005) Understanding the Cyber Defender: A Cognitive Task Analysis of Information Assurance Analysts. Report no. CSA-CTA-1-1. Secure Decisions. Funded by ARDA and DOD.

7 D'Amico, A. and Whitley, K. (2008) The Real Work of Computer Network Defense Analysts. VizSEC 2007: Proceedings of the Workshop on Visualization for Computer Security. Berlin; Heidelberg: Springer-Verlag, pp. 19–37.

8 Foresti, S. and Agutter, J. (2004) Cognitive Task Analysis Report, University of Utah, CROMDI. Funded by ARDA and DOD.

9 Schiffman, M., Eschelbeck, G., Ahmad, D., Wright, A., and Romanosky, S. (2004) CVSS: A common vulnerability scoring system, National Infrastructure Advisory Council (NIAC).

10 Teerlink, S. and Erbacher, R.F. (2006) Improving the computer forensic analysis process through visualization. *Communications of the ACM* 49(2): 71–75.

11 Leung, Y.K. and Apperley, M.D. (1994) A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction* 1(2): 126–160.

12 Merkl, D. and Rauber, A. (1997) Alternative ways for cluster visualization in self-organizing maps. Workshop on SOM '97, pp. 106–111.

13 Baldonado, M.Q.W., Woodruff, A., and Kuchinsky, A. (2000) Guidelines for using multiple views in information visualization. Proceedings of the Working Conference on Advanced Visual Interfaces; May, pp. 110–119.

14 Goldstein, J. and Roth, S.F. (1994) Using aggregation and dynamic queries for exploring large data sets. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence, Boston, MA: ACM Press, pp. 23–29.

15 Boyd, J. (1981) Patterns of Conflict, presentation, http://www.d-n-i.net/boyd/pdf/poc.pdf.

16 http://en.wikipedia.org/wiki/OODA_Loop.

17 di Battista, G., Eades, P., Tamassia, R. and Tollis, I.G. (1999) *Graph Drawing: Algorithms for the Visualization of Graphs*. Englewood Cliffs, NJ: Prentice Hall.

18 Pang, A., Wittenbrink, C. and Lodha, S. (1997) Approaches to uncertainty visualization. *The Visual Computer* 13(8): 370–390.

19 Mahoney, M.V. and Chan, P.K. (2003) An analysis of the 1999 DARPA/Lincoln laboratory evaluation data for network anomaly detection. In: Vigna, G., Kruegel, C. and Jonsson, E. (eds). Proceedings of the Sixth International Symposium on Recent Advances in Intrusion Detection, Lecture Notes in Computer Science, Vol. 2820. Pittsburgh, PA: Springer Verlag, pp. 220–237.