Information Assurance Visualizations for Specific Stages of Situational Awareness and Intended Uses: Lessons Learned

Anita D'Amico and Michael Kocka

Secure Decisions, a division of Applied Visions, Inc.[1]

ABSTRACT

Information visualization has proven to be a valuable tool for working more effectively with complex data and maintaining situational awareness in demanding operational domains. Unfortunately, many applications of visualization technology fall short of expectations because the technology is used inappropriately: the wrong tool applied in the wrong way. A study of visualization techniques as applied to one particularly demanding area – information assurance – leads to the conclusion that there is a proper and formal way to approach designing visualization techniques for maintaining situational awareness in complex domains. Visualization techniques should be specifically designed or selected to align with one of the three identified stages of situational awareness – perception, comprehension, or projection – and with one of five standard uses of visualization: monitoring, inspecting, exploring, forecasting, or communicating. Greater value can be realized by selecting the right visualization technique to focus on each operational task, rather than searching for a single all-encompassing solution to fit every need. Examples of how visualizations can be used to support specific tasks of IA analysis are presented, with examples based on a review of available literature, a formal cognitive task analysis performed by the authors, and lessons learned from direct experience with developing IA visualizations and training analysts in their use.

CR Categories: H.5.2 [User Interfaces]: Graphical User Interfaces, User-Centered Design, Theory and Methods; H.1.2 [User/Machine Systems]: Human Information Processing; K.6.5 [Security and Protection]: Unauthorized Access; D.2.10 [Design]: Methodologies.

Keywords: Visualization, information security, information assurance, IA, data representation, graphics, computer network defense, Human Factors, Security.

1. INTRODUCTION

Visualizations designed specifically for Information Assurance (IA) emerged as a separate specialty in 1999 when NSA's Office of Research and Technology Application, in conjunction with the General Counsel for Technology, licensed software source code to Raytheon Systems for the development of a product called "Silent Runner." That same year the SecureScope™ system for visualizing IA events emerged as a viable prototype from the US Air Force Research Laboratory, and became a commercially-available visualization tool through DARPA support. Other IA visualizations, such as Renoir, were designed specifically for use by the US government during this period. In 2001 commercial security information management (SIM) products such as Intellitactics began incorporating visualizations, e.g. those provided by Advizor [1], into their central management consoles to facilitate event correlation.

Over the past six years many lessons have been learned about the content and form of IA visualizations. In 2004 Secure Decisions, under the sponsorship of the intelligence community, reviewed the literature on best practices in visualization as applied to IA [2] and conducted a cognitive task analysis (CTA) of 41 IA analysts [3]. We have gained considerable insight into how to design visualizations to support IA analysis through the literature review, the CTA, and through our own direct experience with developing IA visualization tools and training IA analysts in the use of those tools.

The goal of this paper is to provide insight and specific examples of how IA visualizations can be used for specific purposes: namely, to enhance each of the three stages of situational awareness described in the literature [4], and to support IA analysts in their use of visualizations for monitoring, inspecting, exploring, forecasting, and communicating IA information. We use the term "information assurance" (IA) which is used by the US government, rather than the term "information security" which is used in the commercial world, because most of the experiences of the authors and examples cited have been drawn from our extensive experience in providing visualizations for IA analysts in the US government. The subtle distinctions between the terms information assurance, information security, and computer network defense (CND) are not relevant to this discussion of visualizations. The information offered herein is applicable to all of those highly related areas.

2. USING VISUALIZATION TO PROMOTE IA SITUATIONAL AWARENESS

IA analysts strive to attain and maintain Situational Awareness (SA) of cyber threats to their networks. Visualizations are one tool they may use to enhance their situational awareness. How operators achieve situational awareness has been studied in many other domains, and research on situational awareness has been published in the psychological and human factors literature over the past 15 years [4] [5] [6]. Mica Endsley, one of the most widely published scientists in situational awareness, describes situational awareness as simply "knowing what is going on around you" and, within that knowledge of your surroundings, knowing what is important. One doesn't need to know *everything*, only those

[1]6 Bayview Avenue, Northport, NY 11768
AnitaD@SecureDecisions.com
This work was supported in part by Air Force Research Laboratory and Advanced Research and Development Activity (ARDA) Contract F30602-03-C-0260.

Workshop on Visualization for Computer Security
October 26, Minneapolis, MN, USA
0-7803-9477-1/05/$20.00 ©2005 IEEE.

things necessary to make good decisions within the timeframe in which they must be made.

Situational awareness is not a simple, atomic state: it is a process wherein one's perspective changes based on how things are looked at, what data is available, or what particular goal one is trying to achieve. This process of situational awareness can be viewed as breaking down into three major stages: perception, comprehension, and projection:

- *Perception* refers to knowledge of the elements in the environment that one must know about, such as knowing what the Intrusion Detection System (IDS) alerts are, as well as the time they occurred and what sensors are reporting them.
- *Comprehension* refers to how people combine and integrate the elements they perceive, to derive meaning from them with respect to their goals; figuring out, for example, from a deluge of alerts that suspicious activity on a specific mission-critical database server or an e-mail exchange server requires greater attention than similar activity on less-essential assets. Comprehension is, in essence, knowing when you have perceived something important.
- *Projection* is the individual's ability to project forward in time to anticipate future events. For example, mentally calculating that if the current sequence of suspicious events continues, and they are coming from the same source, then the next likely event will be of a specific type within an estimated timeframe. Projection helps one to decide on the next course of action.

Visualization and user interface techniques should be selected based upon which of these three stages of situational awareness one wishes to enhance.

3. Improving Perception with Visualization

To *perceive* what is happening, IA analysts need to find the "signal in the noise". Visualization can be used to highlight that signal to help it stand out from the noise.

In many cases the data associated with the attacker is intermingled with a substantial amount of other data. The IA analyst needs assistance in finding the relevant information amidst the irrelevant data. For example, the analyst may be looking for any one source IP address that is scanning many destination IPs within the same site's network traffic. He may or may not know the particular source IP address he is looking for, and the data related to one source IP scanning many destination IPs is embedded in *millions* of transactions occurring over a full day. In this situation, the role of visualization and its associated user interface is to *highlight the signal in the noise*: to provide the analyst with a method for seeing the one-to-many relationships, making them stand out from the sea of unrelated data.

Traditional methods such as scatter plots, combined with easy-to-use filtering capabilities, can be applied to this problem to enhance the perception phase of situational awareness. This is illustrated in the following sequence of scatter plots of connections between Source IP addresses and Destination IP addresses.

The analyst would start out with considerable noise, as seen in here in Figure 1:

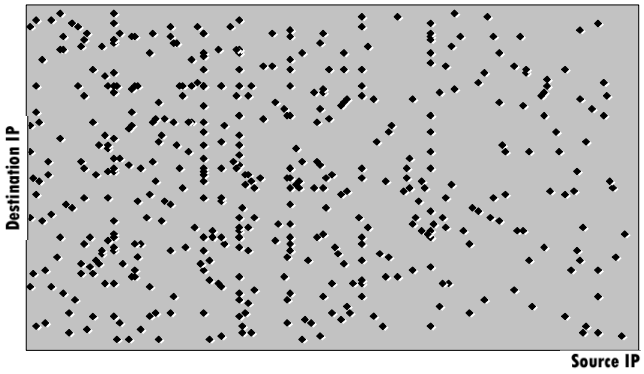

Figure 1: Noisy scatter plot of source and Destination IP addresses

We provided the analyst with the ability to filter out those connections in which the Destination IP did not return bytes to the Source IP, based on the assumption that a scan that returned no data can be considered benign. Figure 2 shows the same scatter plot after filtering out such scans, where the Destination IPs did not return bytes to the Source IPs. Patterns of dots forming vertical lines emerge, indicate the Source IPs that are scanning the military network.

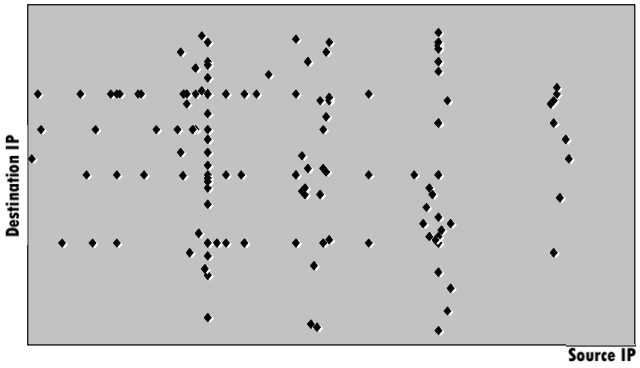

Figure 2: Scatter plot showing only Destination IPs to return bytes

In Figure 3 we further filter the plot to eliminate all .mil to .mil traffic, on the assumption that traffic within the protected .mil domain is secure and can be ignored. What remains are connections where military Destination IPs have returned bytes to non-military Source IPs: something an IA analyst would definitely want to investigate further. These suspicious connections stand out as patterns of horizontal lines associated with specific Destination IPs, as shown in Figure 3:

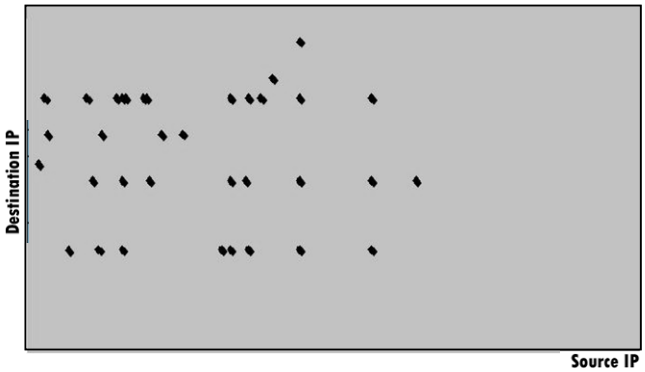

Figure 3: Scatter plot excluding .mil-to-.mil connections

4. Enhancing Comprehension through Visualization

During the *comprehension* phase of situational awareness, IA analysts put the pieces of the puzzle together to create a mental model of what type of attack is underway, including who might be responsible for it, and the existing footprint of the attacker in the network. He or she associates a series of activities into a cohesive, related group.

Visual data presentation can facilitate the rapid comprehension of a sequence of interconnected events. For example, when an analyst explores data looking for a path that an attacker may have taken through the network, he tries to ascertain what systems the attacker may have come in contact with, and possibly exploited, along the way. He may sort through data to discover that Source A connected to Destinations B, D, G and V. He may then look at B, D, G and V to determine with whom they connected, after their contact with Source A, and whether any of those connections were atypical. In the process of doing this the analyst is mentally constructing a network of suspicious transactions.

In this case, visual representations of the relationships that the analyst uncovers can assist him in seeing possible routes the attacker may have taken, and in communicating the sequence of attacker's actions to others. A link analysis visualization of the connections between various entities, and an animation of a possible path an attacker could have taken, can help the analyst gain insight into the attacker's activities.

5. Enhancing Projection through Visualization

In the *projection* phase of situational awareness, IA analysts project into the future to hypothesize what future actions an attacker could take if allowed to roam through the network, and what the effects might be on the network if the attacker's IP address or port of entry is blocked. Forecasting future threats and exploits is also part of this projection phase.

IA analysts engaged in projection rely on visualizations that use timeline analysis and link analysis, such as those offered by I2's Analyst Notebook [7], in this phase of their analysis.

6. What IA analysts do, and how they use visualizatons

Designers of IA visualizations and associated user interfaces need to know what analysts do, and how they intend to use visualizations, before beginning the visualization design process.

6.1 Analyst Job Functions

Analysts have job functions that are either reactive or proactive [8]. *Reactive activities* are described by Carnegie Melon University (CMU) as triggered by an event or request, such as a report of a compromised host, wide-spreading malicious code, or something that was identified by an intrusion detection or network logging system. *Proactive activities* "provide assistance and information to help prepare, protect, and secure constituent systems in anticipation of future attacks, problems, or events." In addition, there are functions that are neither proactive nor reactive, which CMU researchers refer to as *security quality management activities*: these are services that support information security but aren't directly related to a specific security event, such as product evaluations, training, and disaster recovery planning.

Knowing whether the visualization user is working in reactive or proactive mode has direct implications for the types of data one visualizes, the level of detail to which it is presented, and the requirements for rapid updating of the visualizations.

Most IA analysts work in reactive mode. They look at network traffic and other data to draw conclusions about whether the information assets they are responsible for protecting are being attacked, the nature of the attack, its origin, what the attacker might do next, how the attack might impact the organization, what courses of action are available to defend against the attack, and how such an attack might be avoided in the future. To answer these questions IA analysts consult the output of automated systems that provide them with network data, which has been automatically collected and filtered to focus the analyst's attention on data most likely to contain clues regarding attacks.

When working in reactive mode, analysts need to visualize data from their "sensors." The form of the data varies and may include IDS logs, vulnerability scanner event logs, and TCPDUMP data. Analysts working in reactive mode also have varying time constraints.

The "real time" analyst may have as little as 90 seconds to make a decision regarding whether activity is suspicious or not. To support real time analysis, visualizations must automatically update with new data. Visual attributes should clearly distinguish new data from old data, and highlight data of high priority. Mechanisms for doing this can be fairly simple, such as tables that are sorted by timestamp and use color codes to identify priority levels.

Other IA analysts spend hours looking for patterns in current and past data. They look for patterns related to the time of suspicious activities, IP addresses that are the source or destination of suspicious activity, destination ports, and any other patterns that can help them to detect attacks and profile attackers. Analysts engaged in reactive pattern analysis will typically want to visualize data extracted from a long time period, such as several days or weeks. Automated update of visualizations is not required and can, in fact, be disconcerting if the analyst is exploring and studying a dataset for patterns. Visualization techniques that summarize data and present it on multiple axes are needed for pattern analysis.

After examining the patterns in data from suspicious activities, analysts may engage in some proactive analysis in which they postulate what the next action of an attacker might be, given the historical pattern of prior attacks. Analysts may also engage in proactive threat analysis, in which they identify potential attackers or attack groups that have not yet been detected on the defended network, but are expected to attack in the future. Here techniques such as link analyses – that connect events or suspects to each other – are useful. Time lapse replays are also useful techniques that move the viewpoint along a time line and allow the user to mentally project the timeline forward.

6.2 Categories of Uses of Visualization

In an extensive literature review that Secure Decisions conducted on best practices in IA visualization [2], we found that there are certain universal applications of visualization. Irrespective of the domain – it could be IA, weather forecasting, intelligence analysis, or financial analysis – there are five major ways that people use visual data presentation: monitoring, inspecting, exploring, forecasting, and communicating.

A given visualization may be complete and appropriate for one purpose, yet be lacking or misleading for another. This situational dependence suggests that no single data presentation design may be generically suitable and, in fact, reveals that attempts to create such a "universal" visualization to address a user's needs across all of these potential uses will be misguided. Users are better served by a "tool box" of different visualizations tailored to the specific requirements of each category of use of visualization.

There is not a simple one-to-one relationship between these five categories of uses of visualization and the three stages of situational awareness described earlier. Neither is there a direct relationship to the IA analysis activities for which they are used. Rather, there are overlaps that must be understood and kept in mind when designing visual interfaces to meet the needs of a particular task. The diagram in Figure 4 illustrates this landscape:

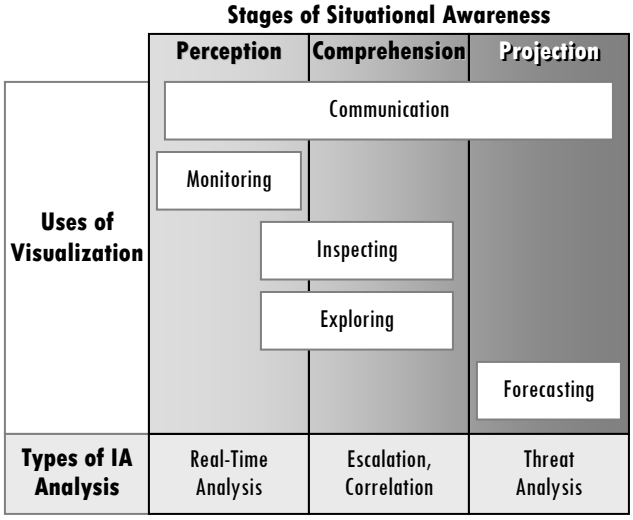

Figure 4: Relationship between the stages of situational awareness, the uses of visualization, and the types of IA analysis performed using those types of visualization

The remainder of this section explores these five different uses of visualization, and how they relate to the three stages of situational awareness. From this discussion it will become apparent how the specific needs of each potential can direct the designer to create visualization techniques that are properly focused on those needs, and thereby more effective for the user.

6.2.1 Monitoring

Someone who is *monitoring* a system is watching an ongoing phenomenon in which data may be continually changing. It is part of the perception stage of situational awareness. Visualizations that facilitate monitoring must be able to be updated with the newest data, have unambiguous indicators of activity, present information within a frame of reference such as a status between "normal" and "dangerous", cover the broad area being monitored rather than partial views, and provide summary status. The user needs to comprehend the status of the system at a glance, and to perceive state changes. The form of presentation may be as varied as the situation being monitored: tables, pie charts, or "stoplight" graphics that provide red/yellow/green indicators of general state are often used in monitoring tasks. "Digital dashboards" are a common embodiment of monitoring, combining several charts, plots, or gauges to provide an overview of organizational performance [9].

IA analysts monitor individual networks or groups of networks to perceive the general security status and gross measures of deviations from normal. (Monitoring is distinct from "inspecting" for specific suspicious activities, which we discuss later.) When monitoring a site, the analyst needs visual aids that represent deviations from normal in activities such as:

- Number of connections – The absolute number of connections for each node on the network is a bellwether. Graphical depictions of normal and current connections per node, per site, and per subnet are useful to general monitoring.
- Amount of data transferred – Data exfiltration is a major concern to many organizations, so when large data transfers occur from a location that is not known for such activity the analyst's internal alarms go off. Visual indicators of data transfer sizes could be used to identify such problems.
- Number of overall security events – When several IDSs are all firing heavily, the analyst looks for a common source. Likewise, a particularly quiet sensor could indicate that a sensor is not functioning properly either because of maintenance issues or perhaps exploitation by an attacker.

Managed Security Service Providers (MSSPs) typically implement a "big board" display where they monitor the status of many of their accounts simultaneously. By watching the number of "tickets" they have to respond to for each customer, the MSSP can tell whether this is a normal, light, or exceptionally heavy day. They can also see general patterns across their customer base, such as whether there are more suspicious activities than usual across a broad base of their customers. Monitoring of general status across several customer groups can help analysts to identify a Zero Day Attack. Color-coded tables, geographic maps with "hot spots" highlighted, and histograms and bar charts are useful methods for presenting these big board pictures. The SANS Internet Storm center at isc.sans.org is illustrative of these techniques; a sample from their web site is captured in Figure 5. Another noteworthy example is MITRE's IWViz, which depicts volumes of alerts overlaid on a geographic map.

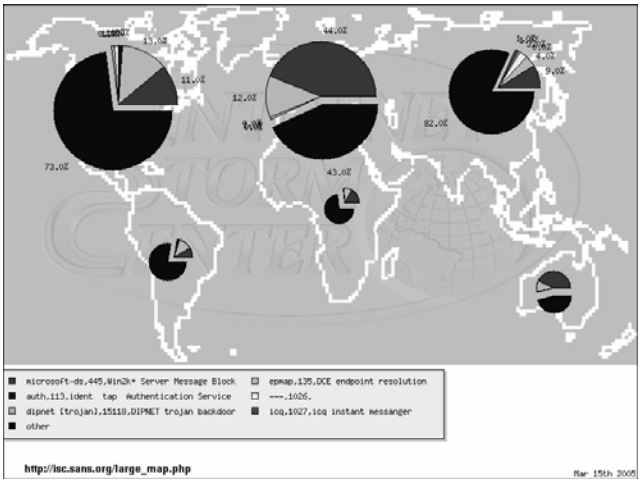

Figure 5: Global monitoring visualization, from SANS *Internet Storm Center*

6.2.2 Inspecting

During the *inspection* process an analyst searches for specific details, requests clarification, and finds data to test hypotheses. Inspection is part of the perception stage of situational awareness, and may continue into the comprehension stage of situational awareness. Inspecting is often associated with analytic discoveries, where the analyst has reached an understanding of the situation by examining and associating pieces of information. Inspecting may be considered top-down, goal-based, or task-based. Operators may or may not be expert, but will likely be heavily influenced by their mental model of the data domain and the system [10].

Most analysts who engage in intrusion detection are engaging in some type of inspection. During the process of inspection, the analyst constrains the dataset that he or she is focused on through judicious filtering, such as illustrated earlier in Figure 1 through Figure 3.

Visualizations may facilitate or enhance the inspection process through representation of:

- Many-to-One Connections – Reactive IA analysts try to find many Source IPs that are attempting to connect to one particular Destination IP. This could be a sign of a coordinated attack on a particular target, or perhaps of a Denial of Service attack. Link analysis is an example of a visualization that can be used to depict such connection patterns.
- One-to-Many Connections – One can infer that a scan is occurring by viewing one Source IP attempting to communicate with many Destination IPs. Visual representations of this phenomenon will show a fan pattern; this visual representation is so strong that analysts refer to "seeing the fan" when discussing one-to-many connections.
- Number of Connections – The same type of visualizations as described under monitoring can be applied to the inspection task, however the dataset under study is typically more constrained.
- Amount of Data Transferred – In the inspection process the analyst focuses on the data transfers of selected hosts that are under study, rather than on general trends in data transfer. To represent data transfer volume one can use line thickness between two IP addresses and the length of a bar under user-selected hosts.
- Length of a Connection – The length of a connection between two nodes can be represented by bar heights. In addition, highlighting or blinking can be used to notify the analyst that the connection length exceeds a certain threshold.

6.2.3 Exploring

Exploration is characterized by undirected perusal, opportunistic discovery without *a priori* clues, novel data combinations, interactive experimentation with data views, finding data regions of interest for analysis, and hypothesis generation. Exploration relates to the perception phase of situational awareness when the analyst is striving to see patterns, and relates to the comprehension phase when he or she begins to explain the findings and assess the situation. Exploring may be considered bottom-up and data-driven, where a variety of options are generated and evaluated but only a few are finally selected as interesting. As with inspecting, analysts will be heavily influenced by their mental model of the data domain and the system [10].

In our experiences with IA analysts we found that exploring is less likely to be performed by real time analysts, who are constrained by time, and more likely to be performed by analysts dedicated to correlation. These analysts search through a day's or week's worth of data, often across many sites, looking for unusual trends to "pop out at them". In fact, the popping-out that occurs is actually a cognitive event, when the analyst associates several pieces of information with each other and adds a hypothesis for why these events are all related. Data visualization enables such *ad hoc* "visual discovery" and recognition of patterns, trends, and anomalies.

Visual data presentation can be very useful for these *ad hoc* types of exploration, as certain patterns are easily comprehended when presented graphically. A so-called "draftsman's plot" is an example of a helpful presentation: it is a matrix of scatter plots that permute the variables mapped to the scatter plot axes. This permits seeing many characteristic distributions side-by-side. Visualizations can reveal distinctive patterns: time trends emerge when frequencies over time are presented as bar graphs, geographical concentrations are easily perceived when events are placed on a map, and paths that are traversed by attackers can be seen more easily when their path is represented as movement through a network topology.

In Figure 6 the SecureScope time wall visualization is used to depict time patterns in security events [11] [12]. The analyst has explored the data and discovered that the critical mail server in the New York IT Support group has been the victim of a series of security events, starting with unusual password activity on Sunday and progressing to a denial of service event on Friday. This visualization supports exploration (as well as forecasting, which is discussed below) in that it provides historical context from which predictions can be made.

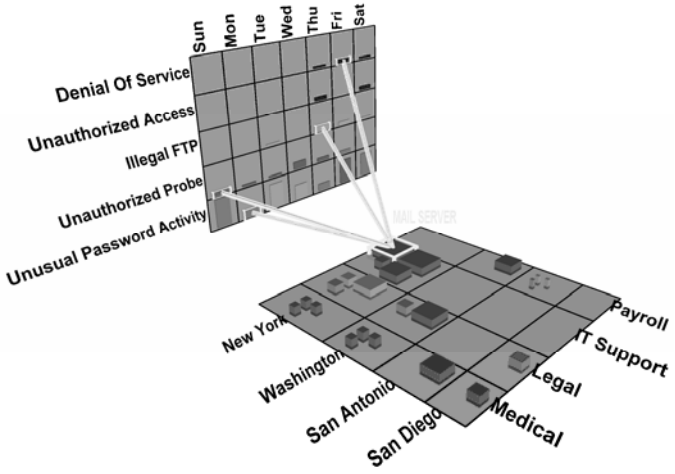

Figure 6: Time patterns depicted in a SecureScope visualization

6.2.4 Forecasting

The goal of *forecasting* can be to either find the likely future state presuming the current progression continues without intervention, or to determine a particular future state based on potential courses of action. Absent a predictive model, forecasting is driven by pattern matching and trending. It uses extrapolation and correlation to extend situational awareness from comprehension forward into the projection phase. IA analysts forecast the attacker's actions, and the impact of an attack if left unchecked, and forecast new exploits that they expect to hit their network in upcoming days or months.

Forecasting requires data that details the current state, and a model (even if only an implicit mental model) describing system behavior. A forecast is often achieved by matching the current situation against the past, and projecting the future based on past progressions. To support forecasting, interfaces should support ready comparisons, pattern-based searching, and trending. Visualizations can help the analyst to forecast by showing the historical activity of the attacker, from which the analyst can infer the next likely actions. Notably, event sequences (that preserve ordering but remove timing) may be better than quantitative timelines for locating precedents.

Forecasting can also be facilitated by visualizations that represent a sequence of actions against a background that combines time, location, and/or organizational structure. By animating and replaying the visual representation, the analyst can see progression and infer the speed and direction of the next action, and even predict who the next victim might be, based on where the attacker has already been, when he was there, and what types of targets in the organization he attacked.

6.2.5 Communicating

Visual data presentation is a useful means for *communicating* with other people, reporting to them, and educating them about one's activities. Educating may be associated with training others to perform particular tasks, teaching others about particular con-

cepts, communicating findings to colleagues or laypersons, and/or documenting decisions for review or justification.

Visual data presentations that support communication of analysts to superiors and subordinates can help communicate critical information in a manner that is easy to comprehend by those not directly engaged in the analysis. They can also help analysts to explain why they formed certain hypotheses or took certain actions, by presenting knowledge that may not be available to all concerned. For example, if those to whom the IA analyst is communicating are unfamiliar with the network topology, but an understanding of the topology was a critical component of the analyst's decision making, then a visual depiction of parts of the network topology can provide the audience with the basic knowledge they need to understand the analyst's actions.

IA analysts regularly engage in educating others or communicating to them the results of what they found in their analyses. There are many forums for communication and education, such as official reports that include graphical depictions of events and statistics, daily PowerPoint briefings, or an electronic bulletin board shared by fellow analysts. Each of these forms imposes its own demands on visualizations and their associated user interfaces. For example, official reports and daily briefings require that visualizations be captured as screen shots and dropped into Microsoft Word or PowerPoint documents, while electronic bulletin boards impose design requirements to share visualizations and interact with them through a web portal.

7. SUMMARY

There is no silver bullet in IA visualizations; one single visualization technique will not meet all of the needs of IA analysis. Visualization techniques should be selected to align with one of the three stages of situational awareness: perception, comprehension and projection. In addition, the designer of IA visualizations and user interfaces must consider the intended purpose of the visualizations: monitoring, inspecting, exploring, forecasting or communicating.

8. ACKNOWLEDGEMENTS

The authors wish to acknowledge Daniel Tesone and Brianne O'Brien, both of Secure Decisions, for their work in performing the cognitive task analyses and in training IA analysts in visualization tools, both of which form the basis of this paper. We also wish to acknowledge Mr. Richard Brackney of the Advanced Research and Development Activity (ARDA) and Dr. Kirsten Whitley of DOD for their funding of the literature review of best practices in IA visualization under Contract F30602-03-C-0260, and the Air Force Research Laboratory as the contracting agency.

REFERENCES

[1] http://www.advizorsolutions.com

[2] Kocka, M. and D'Amico, A. (2004), *Best Practices for Data Presentation*, Report CSA-VIZ-1 under Contract No. F30602-03-C-0260 issued by USAF, AFMC Air Force Research Laboratory

[3] D'Amico, A., Tesone, D., Whitley, K., O'Brien, B., Smith, M. and Roth, E. (2005), *Understanding the Cyber Defender: A Cognitive Task Analysis of Information Assurance Analysts*, Report CSA-CTA-1-1 under Contract No. F30602-03-C-0260 issued by USAF, AFMC Air Force Research Laboratory

[4] Endsley, Mica R. (1995), *Toward a Theory of Situation Awareness in Dynamic Systems*, Human Factors, 37(1) 32-64.

[5] Adams, M.J., Tenney, Y.J., & Pew, R.W. (1995), *Situation awareness and cognitive management of complex systems*, Human Factors, 37 (1).

[6] Endsley, M.R. & Garland, D.J (eds.) (2000), *Situation Awareness Analysis and Measurement*, Mahwah, NJ: Lawrence Erlbaum Associates.

[7] http://www.i2.co.uk/Products/Analysts_Notebook/default.asp

[8] Killcrece, G., Kossakowski K.P., Ruefle, R., & Zajicek, M. (2003), *State of the Practice of Computer Security Incident Response Teams (CSIRTS)*, Technical Report CMU/SEI-2003-TR-001, ESC-TR-2003-001.

[9] P. Russom (2002), *Trends in Data Visualization Software for Business Users*, DM Review.

[10] Endsley, M.R., Bolte, B., & Jones, D. (2003), *Designing for Situation Awareness: An Approach to User-Centered Design*, New York, NY: Taylor and Francis. pp. 13-18

[11] http://www.securedecisions.com

[12] D'Amico, A. & Larkin, M. (2001), *Methods of Visualizing Temporal Patterns in and Mission Impact of Computer Security Breaches*, Proceedings of the DARPA Information Survivability Conference and Exposition II, IEEE Computer Society, Los Alamitos, CA.