Visual Discovery in Computer Network Defense

Anita D. D'Amico, John R. Goodall, Daniel R. Tesone, and Jason K. Kopylec
Applied Visions

Computer network defense (CND) requires analysts to detect both known and novel forms of attacks in massive volumes of network data. VIAssist is a visualization framework based on a comprehensive cognitive task analysis of CND analysts, and so fits their work practices and operational environments.

The systematic analysis of computer network activity is a field of expertise that has grown in response to the exponential increase in computer viruses, network intrusions, and denial-of-service attacks on computing infrastructures. Government agencies, managed security service providers (MSSPs), and corporations have established network or security operational centers (NOCs or SOCs) staffed by analysts whose job is to detect and defend against security breaches in the critical information infrastructure. These analysts comb through large volumes of network data and intrusion detection alerts to discover real attacks amidst the preponderance of false alarms, and they review massive amounts of packet data to identify suspicious activities that might have slipped past the security sensors placed on or outside the organization's network. These analysts also detect and report a network's vulnerabilities to potential attack, and identify unauthorized usage and policy violations that could expose a network to greater risk of compromise. Some analysts specialize in forecasting new threats and predicting the activities of attackers under observation. Still others engage in forensic analysis of cyberattacks.

The analysts' domain of expertise is referred to as information/network/computer security, or InfoSec in the commercial world and information assurance (IA) or computer network defense (CND) in government. (We'll use "CND analyst" in this article.) Regardless of their job title or subspecialty, these cyberdefense analysts strive to attain and maintain situational awareness regarding the networks they defend and the attackers they defend against. It's through discovering the unexpected that CND analysts detect new versions of malware (such as viruses and Trojan horses) that have passed through their antivirus products, new methods of intrusion that have breached their firewalls and intrusion detection systems (IDSs), and new groups of cybercriminals pressing the attack.

Visualizations that depict patterns in massive amounts of data, and methods for interacting with those visualizations can help analysts prepare for unforeseen events. To help us design visual tools for CND analysts, we conducted a cognitive task analysis (CTA; see the "What Is Cognitive Task Analysis?" sidebar) to gain a deeper understanding of how they perform their craft: their cognitive needs, how they currently use visual data presentation methods in their analysis, and how visualization technologies could enhance their situational awareness in the future.[1] Findings from these observations led to the design of the Visual Assistant for Information Assurance Analysis (VIAssist).

Computer network defense: The analytical process

Most CND analysts work reactively. They look at network traffic and other data, asking questions and drawing conclusions about whether the information infrastructure they're protecting has been attacked, the nature of the attack, the attacker's identity, and the required response. To do this, CND analysts consult automated systems' output, filtered to focus their attention on data most likely to contain clues regarding attacks. Several types of automated systems—IDSs, vulnerability scanners, antivirus systems, and system administration tools—produce log files that an analyst can inspect to detect suspicious activities.

IDSs are the most frequently used of these systems. An IDS is a sensor that monitors network traffic looking for suspicious activity. The IDS might be designed to log an alert when it detects a particular signature—a sequence of keywords, for example—in an incoming packet, or anomalous activities—such as increased activity on a certain port. IDS logs are typically fraught with false positives. An essential part of the reactive analytical process is to determine which of the alerts refer to true malicious activity and which are, in fact, false positives.

Inherent in the CND analytical process is the search for patterns in the data. Analysts look for patterns related to the time of suspicious activities, IP addresses, or ports that are the source or destination of suspicious activity, as well as any other patterns that can help them

 Published by the IEEE Computer Society

detect attacks and profile attackers. After examining these patterns, analysts might engage in some proactive analysis, during which they postulate an attacker's next action, given the historical pattern of prior attacks. They might also engage in proactive threat analysis, in which they identify potential attackers or attack groups that haven't yet been detected on the defended network, but can be expected to attack in the future.

When developing visualizations to support cyberdefense, designers must focus on the target user's specific role. CND roles include triage, escalation, correlation, threat, forensic, and incident response analysis (a full description of these roles is available elsewhere[1]). VIAssist targets analysts who perform escalation, correlation, and threat analysis.

Escalation analysis investigates potential incidents referred from triage, colleagues, and cooperating organizations. It takes hours to weeks, during which the analyst marshals more data, usually from multiple data sources both inside and outside the organization. This data helps the analyst better understand the attacker's modus operandi and goal, and the attack's extent and severity. In the course of the analysis, analysts often discover relationships (such as a common source IP address or similar character strings) between two seemingly disparate security events.

Correlation analysis searches for patterns and trends in current and historical data. It might extend to grouping and investigating related incidents across separate organizations, a task that can take months. Analysts performing correlation analysis engage in extensive information synthesis and data fusion, often without a specific target in mind.

Threat analysis uses additional data sources (such as information from hacker Web sites) to profile attackers and discover their identities and motivations. Analysts in law enforcement and government, and at MSSPs and commercial computer security companies, use this type of intelligence-gathering.

Although these roles are distinct, they share certain characteristics that were important motivators for VIAssist:

- They gather a variety of data into one place, and parse the data in different ways.
- They take time to look at the data and explore it for patterns.
- They try to form relationships across different data sources and security events.
- They capture and share raw and visualized data with others through briefings and email.

Supporting CND through visualization

During our data collection, we observed that in the few instances that visual tools were used, analysts used them almost exclusively for correlation and threat analysis, and for creating postanalysis graphics to be included in informational briefings. We also observed that analysts adapted visualization tools borrowed from non-CND applications to fit their needs, often modifying their data to fit the tool. So, despite a substantial amount of academic work on information security visualization (see the "Visualization for Computer Security" sidebar on the next page for an overview), few CND-specific visualization systems exist in CND operational environments.

We concluded from our data collection effort that improving visualization's effectiveness in CND analysis might not require new visual idioms. Rather, we could modify existing visualization techniques to align with an analyst's role and specific data formats. We could then design a robust user interface framework designed specifically for their work practices, integrating all of the tools they require. We also learned that one type of visualization can't support all the different CND roles, and even a single type of analysis—such as correlation analysis—might benefit from looking at the same data from multiple perspectives.

VIAssist lets CND analysts use a suite of integrated visualizations to examine the same data set from multiple

What Is Cognitive Task Analysis?

Cognitive task analysis is the study of an individual's or team's mental processing, activities, and communications within a specific work context. CTA (and similar methods such as knowledge elicitation, cognitive engineering, and workflow analysis) was introduced as a methodology at about the same time that computer technology was introduced into the workplace (that is, the 1980s) as a mechanism for improving the effectiveness of expert systems that attempted to automate some human decision-making.[1]

CTA elicits information from individuals about the thought processes they use while completing specific tasks. It involves observing individuals as they work and asking directed questions about how they approach problems and decide their next steps, as well as their work's challenging tasks. Because CTA focuses on cognitive processes, rather than the mechanics of completing tasks in the current environment, the results tend to be less affected by tool bias. A CTA's output is a detailed description of the tasks that an individual or team performs, the data on which they operate, the decisions they make, and the processes and activities (cognitive, communicative, and physical) that they engage in to reach those decisions.

During the CTA referred to in this article, we interfaced with more than 40 computer network defense (CND) analysts involved in some aspect of network security. They varied in level of expertise from novice to expert and represented a variety of job titles and work roles.

We used a combination of four knowledge capture techniques: semistructured interviews, observations, review of critical incidents, and hypothetical scenario construction. This last technique involved working with analysts to flesh out an imaginary analysis case, including typical offensive actions taken by a sophisticated attacker and defensive actions by the CND analyst. The exercise revealed the kinds of information analysts seek from available data sources, their knowledge of adversary operations and techniques, and types of connections that they make between seemingly disparate pieces of information.

Reference

1. R.P. Hoffman and D.D. Woods, "Studying Cognitive Systems in Context," *Human Factors*, vol. 42, no. 1, 2000, pp. 1-7.

Visualization for Computer Security

Visualization of computer network traffic typically aims to help users diagnose performance issues or understand communication patterns between nodes, either within a network or in the Internet as a whole. Traditionally, information visualizations of network traffic have used link and node graph-based techniques. For example, SeeNet includes a graph visualization that places nodes according to their natural geographic location and uses thickness and color to encode network statistics to provide a high-level view of network traffic.[1] These kinds of graph-based visualizations are useful in mapping networks and usage patterns by explicitly showing links between nodes, but can have problems with scalability, display clutter, and occlusion. Interaction and focus-plus-context techniques can help deal with some of these problems. These visualization and interaction techniques have greatly influenced more current research specific to computer network defense (CND).

More recently, researchers in academia and industry have focused on network visualizations specifically to address the challenges of CND. It's beyond the scope of this review to discuss all of the research in visualization for computer security. Instead we present an overview of recent advancements in systems that visualize NetFlow data, since our cognitive task analysis revealed analysts' extensive reliance on that data.

Several researchers have used NetFlow data specifically to address security issues. Like VIAssist, NVisionIP visualizes NetFlow data using proven visualization techniques.[2] NVisionIP shows a class-B network's overall state in a scatterplot linked to other views showing increasing levels of detail. IDGraphs attempts to aid analysts in detecting denial-of-service attacks and scans by visualizing NetFlow data as histographs, which map frequency to luminance mapping to merge and summarize graphs.[3] Other systems have visualized network data at lower levels of abstraction, such as packet-level data;[4,5] or at higher levels, such as IDS alerts.[6,7]

References

1. R.A. Becker, S.G. Eick, and A.R. Wilks, "Visualizing Network Data," *IEEE Trans. Visualization and Computer Graphics*, vol. 1, no. 1, 1995, pp. 16-28.
2. K. Lakkaraju, W. Yurcik, and A.J. Lee, "NVisionIP: NetFlow Visualizations of System State for Security Situational Awareness," *Proc. ACM Workshop on Visualization and Data Mining for Computer Security* (VizSEC/DMSEC), ACM Press, 2004, pp. 65-72.
3. P. Ren et al., "IDGraphs: Intrusion Detection and Analysis Using Stream Compositing," *IEEE Computer Graphics and Applications*, vol. 26, no. 2, 2006, pp. 28-39.
4. G. Conti et al., "Visual Exploration of Malicious Network Objects Using Semantic Zoom, Interactive Encoding and Dynamic Queries," *Proc. Int'l Workshop Visualization for Computer Security* (VizSEC), IEEE CS Press, 2005, pp. 83-90.
5. J.R. Goodall et al., "Focusing on Context in Network Traffic Analysis," *IEEE Computer Graphics and Applications*, vol. 26, no. 2, 2006, pp. 72-80.
6. Y. Livnat et al., "A Visualization Paradigm for Network Intrusion Detection," *Proc. IEEE Workshop Information Assurance and Security* (IAW), IEEE Press, 2005, pp. 92-99.
7. K. Abdullah et al., "Visualizing Network Data for Intrusion Detection," *Proc. IEEE Workshop Information Assurance and Security* (IAW), IEEE Press, 2005, pp. 100-108.

perspectives. VIAssist is extensible: by plugging in different tools for different jobs, you can use the same framework to provide visual tools to various analytic roles. The framework also supports multiple data sources. For example, VIAssist can interface to an IDS alert database, network flow (NetFlow) data, or an incident repository. And to maintain high-level context during low-level analysis, we incorporated two-display monitors: one for a summary overview of the network activity, and another for in-depth analysis of a subset of the data, such as a specific security incident (see Figure 1).

Through years of operational testing of CND visualizations, first with our SecureScope system,[2] and more recently with VIAssist, we've also concluded that factors beyond the system's visual aspects determine a visualization system's utility for CND. Visual analytic tools for CND have rarely transitioned successfully from the laboratory into the operational environment because they've failed to fit into the operational workflow, can't keep up with the data volume, don't interface well with legacy systems, or are too complicated to learn. Analysts also need to work with visual representations through media other than a desktop screen. They need to embed them in PowerPoint presentations, print them (often in gray scale), email them to other analysts, or post them on a big-board display—all with as little effort as possible.

We designed VIAssist for use by actual analysts working in operational SOCs with real network data, based on the user needs gathered during the CTA. Our goal was to integrate proven visualization techniques into a robust framework that addresses the real-world requirements of CND analysts working in diverse settings. As part of an iterative design cycle, we showed early prototypes to CND analysts, incorporated their feedback into the system design, and deployed the revised system for further test and evaluation in an operational environment.

CND analysts are busy, and are unlikely to try a new technique unless they're fairly sure it will make a difference in their work. Because we based VIAssist's design on interviews and observations of real analysts and gave them opportunities to comment on mockups and early prototypes, we were invited to test the resulting system in a real operational environment to which few developers of CND visualizations have obtained access.

We also brought the prototype to a military exercise, in which we could interconnect and demonstrate advanced technologies within a realistic environment, and to a CND analyst conference. At the time of this writing, VIAssist is undergoing a six-month evaluation at the Joint Task Force for Global Network Operations, where escalation, correlation, and threat analysts are exercising the system with real-world data.

VIAssist system design

We designed VIAssist to support CND analysts in the following activities:

- discovering new patterns in large volumes of network security data;

- discovering how a specific event is related to other information (for example, specific IP addresses, ports, keywords, and location) available in the massive data set;
- orienting their in-depth analysis of suspicious activity within the context of a network overview that shows areas of unusual activity;
- forming and evaluating hypotheses about a network compromise's existence, nature, source, timing, and extent;
- tracking where they are in the analytical process and how they got there;
- gathering other sources of information (such as IP lookups and hot IP lists) without leaving the visualization console, and incorporating the results into the view; and
- reporting and conferring with other analysts about their findings.

VIAssist can ingest data from various data sources—NetFlows, IDS logs, watch lists of suspicious IP addresses, and more—and show the same data simultaneously in multiple views to reveal patterns that wouldn't have been discovered through a single perspective. VIAssist supports the discovery of high-level patterns and anomalies, while providing context (through the dashboard) for the detailed area of investigatory focus. It provides intuitive visual interaction mechanisms for fast data manipulation, highlighting, and filtering, and for querying outside information sources (such as Whois) and bringing the results into the visualization.

We currently host several proven visualization components to support this visual analysis. Once we have a usable, robust framework design, developing and integrating new visualization components will be a rather straightforward activity.

Finally, VIAssist enables extended analysis by saving and restoring the analyst's place in the discovery process. It also offers collaborative capabilities such as annotating the visualizations and associated data with notes to other analysts, creating and sharing complex criteria expressions, and building briefings directly from the VIAssist console.

Data management

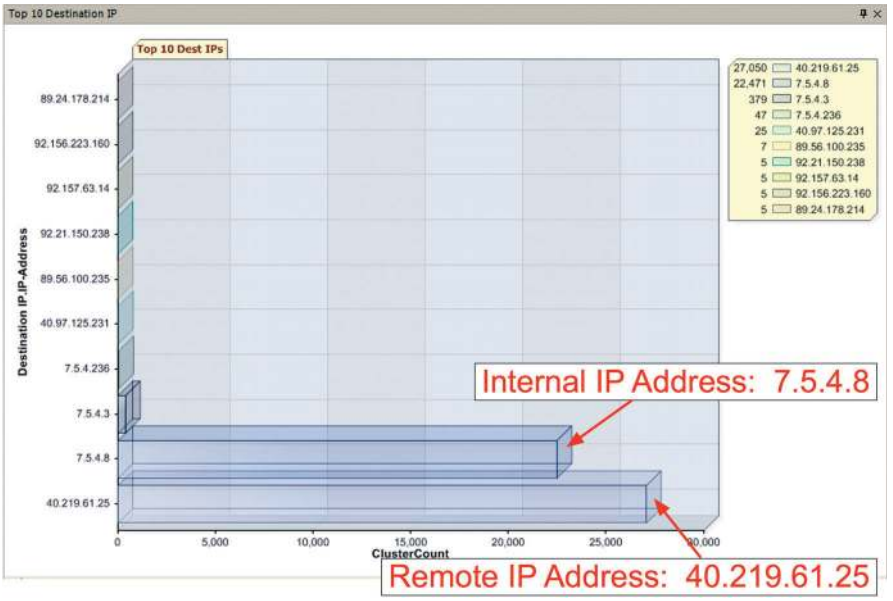
Our CTA found that CND analysts rely heavily on NetFlow data, so this was the first of several data sources that we integrated into VIAssist. NetFlows, which Cisco originally developed to aid network analysts in performance tuning, are uni- or bidirectional aggregations of network traffic between hosts containing various data attributes, such as source and destination address and port, protocol, number of packets and bytes in the flow, and flow start time and duration. We aggregate flows according to TCP sequence number, or a time threshold for UDP communications between unique host pairs.

VIAssist stores this flow data in a relational database. However, rather than require users to query the database using Structured Query Language, like most current CND tools, VIAssist users need have only modest knowledge of SQL. VIAssist emphasizes interactive, visual exploration and discovery. An expression builder—a GUI for building, tagging, combining, and executing expressions—lets even the most novice user create powerful SQL queries using point-and-click. After an analyst creates an expression, he or she can reuse it as criteria for data filtering, display data filtering, or display data highlighting. In addition, VIAssist's various interactive filtering and highlighting controls continually refine the displayed data as the user explores and tests different hypotheses.

1 VIAssist provides one view for in-depth event analysis (on the left monitor), and a dashboard view of global activity (on the right).

The ability to formulate expressions that characterize particular behavior (such as an external IP address connecting to a series of hosts within the same defended subnet, with connection times of a specified length, at certain times of day) has value in both the reactive and proactive analysis modes. Proactive analysis involves anticipating attackers' actions and preparing for them in advance. Analysts formulate hypotheses or hunches about the attack's nature and what the attacker's next moves might be. Tracking these hunches and finding data that matches them is particularly difficult when the data volume is massive and time is limited. However, VIAssist's expression builder lets an analyst formulate a hypothesis about suspicious activity into an expression that will highlight every instance of data meeting its criteria. The analyst can build a library of expressions and permutations of existing expressions, and even share them with other analysts. This catalog of expressions forms a body of knowledge about the types of activities the analyst is anticipating, which the analyst can prune as he or she confirms the hypotheses—and turns them into intrusion detection signatures—or denies them.

VIAssist's smart aggregator helps deal with the typically massive data sets encountered in CND analysis. This tool emphasizes usability by preventing queries that would provide too much data for the visualizations to meaningfully display and that would likely overload the system. For example, if a user asks VIAssist to visualize 30 Gbytes of data with no filtering or aggregation, the application would inevitably grind to a halt because of system load. VIAssist handles this aggregation automatically, attempting to infer the



2 **VIAssist view depicting network traffic volume by destination IP address. The graph clearly shows more activity for two IP addresses: the first, 7.5.4.8, is on the analyst's internal network, and the other, 40.219.61.25, is remote.**

correct level of aggregation to supply for any given request. The idea is to protect users from entering data requests that would overload the system, ensuring that the system is always usable.

Multiple views of the data

Providing multiple simultaneous views of data lets users explore the data from multiple perspectives. Synchronizing those views throughout the analytical process is the challenge. We architected VIAssist to make it appear that all third-party controls belong to the same vendor. Filtering out or zooming in on data in one view automatically updates all other views.

Our visualization framework positions a main view of interest in the display center, with a variable number of supporting views located around it—either on the main or a secondary monitor. We can swap this main view with any of the other views, or relocate views using drag-and-drop.

Along with the investigatory view, a summary dashboard display presents a persistent view of the most critical data attributes, giving a quick overview of the data set that analysts can digest at a glance.

Collaboration

One of the most significant user needs is to work together productively, and to report up the command chain effectively. VIAssist supports this through several mechanisms for collaboration and reporting. These include

- *Sharing items of interest.* Users can set up and share lists of hot IP addresses. For example, if a group of attackers working within a particular IP address space is repeatedly scanning an organization, the organization can globally flag these addresses, so they're highlighted in all users' visualizations.
- *Sharing annotations*. An annotation tool lets users make notes on data and share the annotated data with each other.
- *Communicating hypotheses*. Analysts can use an embedded diary tool, *E-Diary*, to document their working hypotheses without leaving the application. This diary can serve as a shift-changeover communication tool, with each entry tagged with the creator's name.
- *Communicating analytic findings.* VIAssist's report builder lets analysts drag and drop graphical elements from the current display and annotate them to create a PowerPoint or PDF file. Team members can share these files or put them into reports for presentations to management. Templates allow for predefined reports that populate automatically.

These collaboration and reporting tools, although perhaps not as visually interesting as the visualizations themselves, are representative of our approach to meeting users' real-world needs in VIAssist's design.

Discovering the unexpected

CND analysts can use VIAssist visualizations to find interesting patterns in NetFlow or raw network packet data. We present an example of this process at work, based on an attack scenario developed by Skaion Corporation[3] for use in cyberdefense research within the US Department of Defense. The data is synthetic, but the scenario represents one that analysts might face in the wild.

Our analyst, Robin, has been perusing a set of NetFlow data using various charts in the VIAssist dashboard, looking for signs of anything unusual. Some unusual activity, in terms of number of packets in the data set, gets her attention in a bar graph of the top-10 destination IP addresses. On this bar graph, illustrated in Figure 2, the *x*-axis depicts the number of packets, and the *y*-axis depicts the destination IP address.

Robin notes that two hosts, IP 7.5.4.8 and IP 40.219.61.25, clearly have more traffic than all of the others. She knows from experience that 7.5.4.8 is an internal address, and 40.219.61.25 is an external address. (If she was new to the SOC, an embedded lookup tool available in VIAssist could help her determine this.)

Why would the internal host be sending or receiving so many packets? Robin knows it's not one of the few servers that would typically show such a pattern (such as a Web server). It's internal to the network, so it should be trusted. She notes this potential anomaly in her E-Diary.

Of greater concern is the external address, which is receiving or sending many packets to the internal network. This is an unusually high amount of traffic coming from any single untrusted source, which makes it suspicious. Robin knows that this IP address isn't on the SOC's hot-IP list, because it would have been high-

(a) (b)

3 (a) Highlighting clear spikes of network traffic volume in the time histogram simultaneously highlights that data in all other views. (b) Inxight's Table Lens view shows two large clusters for the remote host 40.219.61.25.

lighted as such when she first saw it on the bar graph. She verifies this by consulting the list, which is always available on the dashboard, and finds that no other analyst has put the 40.219.61.25 IP address on the SOC watch list. She adds it to her personal IP list so that the address will be highlighted when she views any other data containing it.

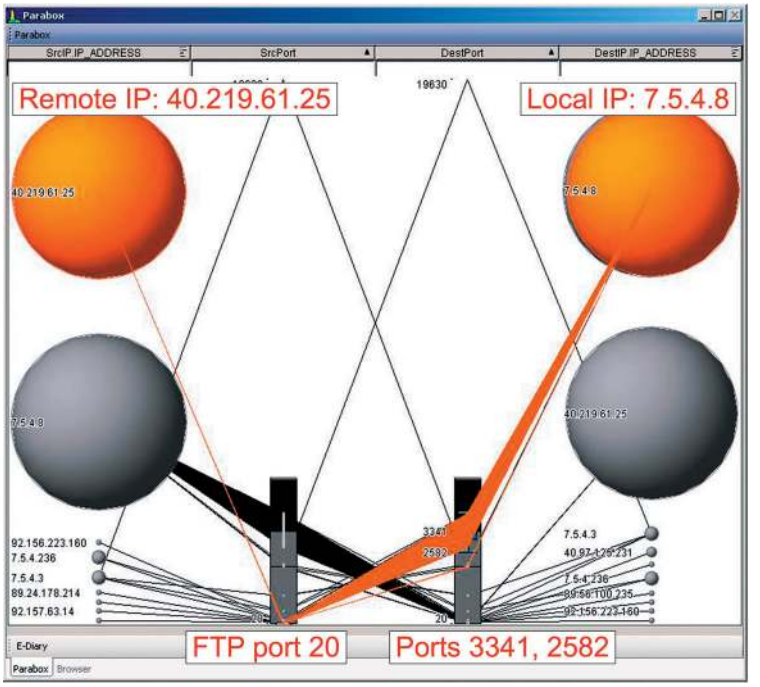
Robin now looks to the rest of the network data in an attempt to discover if this unusual amount of traffic is associated with anything else out of the ordinary. The high rate of activity could just be a remote Web server that the internal address spent a long time surfing. To investigate the issue, she explores the data set using other visualizations.

First, she knows from experience that high traffic activity often occurs at specific times of day (in the morning, for example, when users first log into the network), so she looks at a chart to see if a time pattern is associated with the unusual traffic. As Figure 3a shows, some unusual spikes of activity occur at specific times of the day. She clicks on the spikes and drags the mouse over them to highlight them. Because VIAssist's visualizations are coordinated, the data she selects and highlights in one view is highlighted in all of the views.

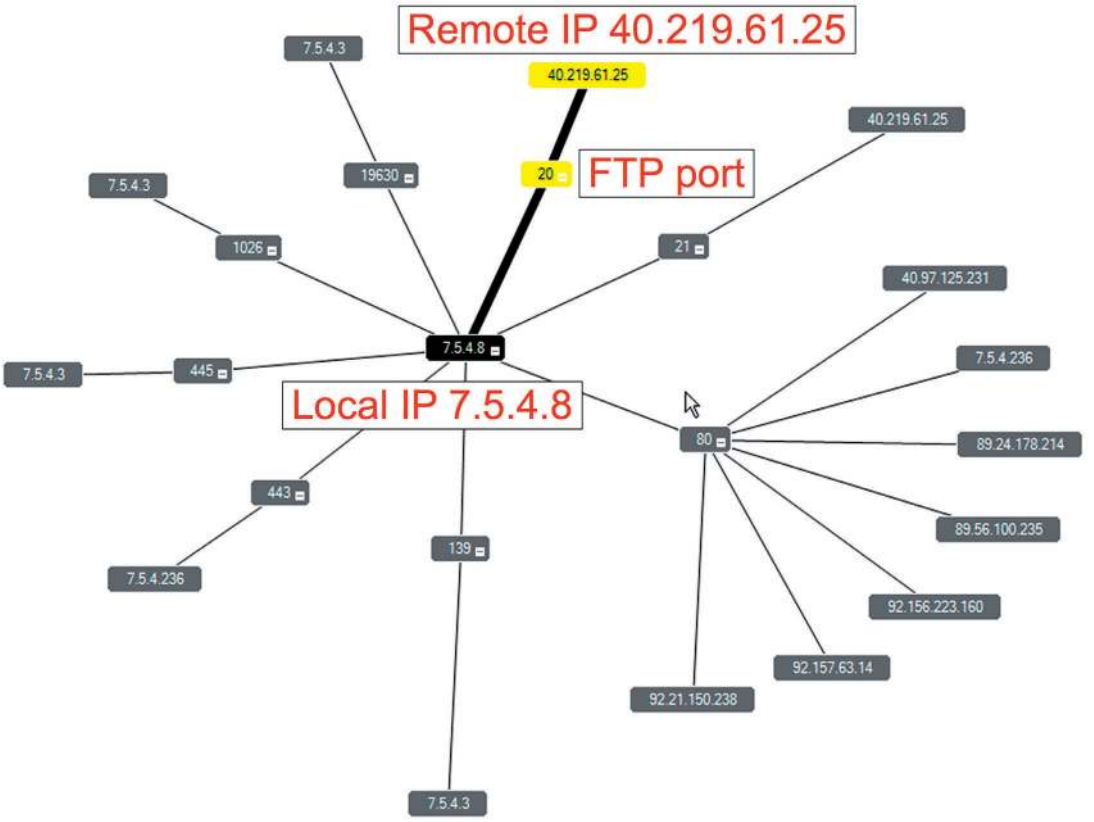
Robin now looks at VIAssist's implementation of Inxight Software's Table Lens view of the same data (Figure 3b). In this view, the source and destination IP addresses, along with the packet count, are displayed, clearly highlighting (in yellow) the three IP addresses contributing to the traffic spikes. She recognizes the single local host (7.5.4.8) and a trusted host (92.156.223.160), both of which are unlikely to be attack sources. However, the remote and therefore the untrusted host (40.219.61.25) has a much larger number of packets going to and from it.

Robin now focuses her analysis on these three IP addresses. She uses a parallel coordinates visualization to determine which ports and protocols the suspicious IP addresses are using. Figure 4 shows VIAssist's implementation of the Advizor Solutions' Parabox parallel coordinates view, in which the bubbles represent (from left to right) source IP, source port, destination port, and destination IP, along the four axes. Bubble size reflects the amount of data involved. The highlighted traffic from the bubble representing remote host 40.219.61.25 to the bubble representing local host 7.5.4.8 shows that all of this traffic passes through FTP data port 20. This bidirectional traffic on the FTP data port indicates that the remote host 40.219.61.25 is acting as an FTP server.

Robin hypothesizes that IP 40.219.61.25 has transferred a large number of files over FTP from IP 7.5.4.8. She finds further support for her hypothesis that inappropriate



4 Network activity of (from left to right) source IP, source port, destination port, and destination IP along the four parallel axes. The highlighted bubbles and links show that all of this traffic goes through FTP port 20.



5 **VIAssist's implementation of Inxight Software's StarTree, showing the local host under investigation at the root (center), with destination port number at the second level and destination IP at the leaves. Line thickness indicates the number of bytes.**

file transfers are occurring when she looks at the patterns of connections between the hosts, using VIAssist's implementation of Inxight Software's StarTree hyperbolic browser display. The StarTree in Figure 5 shows the local host in question at the root of the display, with destination port number as the second level of the hierarchy and destination IP address at the leaves. Line thickness indicates the number of bytes transferred. From this display, it's easy to see that most of the traffic from the local host 7.5.4.8 is going via FTP to the suspicious remote host 40.219.61.25.

Robin must solicit the network administrator's help in determining what types of files are being transferred. She uses VIAssist's report builder to capture the visualizations and prepare a briefing outlining her discovery. When the network administrator receives the report, he captures some of the data being sent to the suspicious host, and finds that it's composed of all of the files in Microsoft Word's "recent files" list. Robin and the network administrator conclude that the user of local host 7.5.4.8 unknowingly let a macro run that copies his recent files, some of which contain confidential data, and sends them to a remote host for some apparently illicit purpose.

To determine whether anyone else has been duped into using this macro, Robin moves 40.219.61.25 from her personal IP list to the SOC's hot-IP list. So, any analyst looking at the data set with VIAssist will see that IP highlighted as a suspicious host to watch and track. As other analysts discover the activities of remote host 40.219.61.25, they annotate the information within VIAssist and share the details using the report builder.

This example is simpler than other hard problems that analysts face, such as detecting a low and slow customized attack on a specific target or identifying the real person (true attribution) behind an attack. The low and slow problem requires the aggregation and systematic filtering of historical data collected over long periods of time—an activity that VIAssist's smart aggregator and filtering mechanisms can support. VIAssist's Whois and country lookup tables can assist with the attribution problem, but attribution poses privacy- and law-related problems that go far beyond VIAssist's capabilities.

Conclusion

As we move forward with VIAssist's development and testing, we'll continue to incorporate user feedback to improve the system's design. Additionally, we'll formally evaluate VIAssist in an experiment that compares performance on a variety of CND tasks using VIAssist and commonly used nonvisual CND tools. This experiment will be an important contribution to the security visualization community, because to date no definitive study objectively measures visualization's contribution to CND effectiveness. Results from prior work on visualization's impact in other domains of expertise vary considerably based on use, users, and circumstances, so they can't be directly applied to CND.

This comparative evaluation will clarify visualization's value in general, and VIAssist's in particular, to information security.

We'll also explore additional collaborative analytic functionality that we can embed in the system—for example, letting analysts more easily share their insights and comment on each other's findings. Communicating analytic findings, which we currently support through our report builder, can be even more robustly supported. For example, VIAssist could help analysts share not only insights gleaned from visual analysis but the path they took to reach a certain conclusion. Collaboration is a capability that visualization rarely enables and indeed, because of most tools' proprietary nature, often hinders. In addition to these kinds of asynchronous collaborative capabilities, we're exploring methods of real-time, collaborative visual analytics that will let analysts interactively explore the same data within a shared workspace. ■

Acknowledgments

The US Department of Defense sponsored VIAssist's development under contract F30602-03-C-0260, with the Air Force Research Laboratory (AFRL) as the contracting agency. We acknowledge the continuous beneficial guidance during VIAssist's development and testing offered by Kirsten Whitley of the US Department of Defense; Walt Tirenin of the AFRL in Rome, New York; and Robert Nine and the J2 staff of the Joint Task Force on Global Network Operations. The views and conclusions contained in this document are those of the authors, and should not be interpreted as representing the official policies, either expressed or implied, of the US government.

VIAssist includes Star Tree and Table Lens visualizations, which are trademarks of Inxight Software (http://www.inxight.com). VIAssist also contains visual charts from Advizor Solutions.

References

1. A. D'Amico et al., "Achieving Cyber Defense Situational Awareness: A Cognitive Task Analysis of Information Assurance Analysts," *Proc. Human Factors and Ergonomics Soc. 49th Ann. Meeting*, HFES Press, 2005, pp. 229-233.
2. A. D'Amico and M. Larkin, "Methods of Visualizing Temporal Patterns in and Mission Impact of Computer Security Breaches," *DARPA Information Survivability Conf. and Exposition* (DISCEX II), IEEE Press, 2001, pp. 343-354.
3. DTO Reference Data Set 4s4, Skaion Corp., North Chelmsford, Mass., 2005.

Anita D. D'Amico *is the director of the Secure Decisions division of Applied Visions. Her research interests include information visualization, cognitive analysis, and transitioning technology into the operational environment. D'Amico received a PhD in psychology from Adelphi University. Contact her at anitad@securedecisions.avi.com.*

John R. Goodall *is a senior analyst with the Secure Decisions division of Applied Visions. His research interests include workplace studies to inform design, information visualization, and usability for computer security. Goodall holds a PhD in information systems from the University of Maryland, Baltimore County. Contact him at johng@securedecisions.avi.com.*

Daniel R. Tesone *is a project engineer with the Secure Decisions division of Applied Visions. His research interests include designing extensible visualization systems that are aimed at enhancing end users' productivity. Tesone has a BS in computer science from the State University of New York at Stony Brook. Contact him at dant@securedecisions.avi.com.*

Jason K. Kopylec *is a senior software engineer with the Secure Decisions division of Applied Visions. His research interests include information visualization, interaction design, and computer network defense. Kopylec has an MS in computer science from Columbia University. Contact him at jasonk@securedecisions.avi.com.*