Data Visualization for Cyber Security

Lane Harrison

Abstract

Data visualization is an indispensable means for analysis and communication, particularly in cyber security. Promising techniques and systems for cyber data visualization have emerged in the past decade, with applications ranging from threat and vulnerability analysis to forensics and network traffic monitoring. We revisit several of these milestones in this chapter.

Beyond recounting the past, however, we uncover and illustrate the emerging themes in new and ongoing cyber data visualization research. We explore the need for principled approaches towards combining the strengths of the human perceptual system with analytical techniques like anomaly detection, for example, as well as the increasingly urgent challenge of combatting suboptimal visualization designs— designs that waste both analyst time and organization resources.

1 Introduction

Unfortunately, cyber security needs data visualization.

Cyber security needs visualization because— in practice— security requires substantial human involvement. Even small organizations require constant time and attention from trained Security Information Workers (SIWs) to ensure acceptable levels of security. SIWs spend much of their time in security operations: scanning devices on their network for vulnerabilities, for example, or analyzing incoming network traffic for malicious activity. Their limited time and attention is fragmented between collecting and analyzing data, and using it to form and prioritize changes in their organization's network and systems. While hardly noticeable at first glance, visual representations of data: bar charts, pie charts, line charts, and the like, are ubiquitous in these operations.

The need for data visualization in cyber security is unfortunate because many or-

ganizations would happily relegate their security operations to intelligent systems, if it were possible.

Intelligent systems operate much faster than human operators, and are less prone to error than we are. The transition to intelligent systems is tempting for many organizations, especially given recent advances in artificial intelligence and machine-learning. Machine learning can now process massive streams of diverse types of data from both inside and outside an organization, for example, providing models that capture malicious behavior. Artificial intelligence, similarly, can analyze network infrastructure to suggest changes that help avoid misconfiguration. These advances represent new ways of thinking in security.

While promising, these advances have not been adopted in operational contexts, nor will they replace the security analyst. Experts in machine learning argue that, even when intelligent systems reach the point of making operational decisions, human judgement will still be necessary for managing the systems themselves.

This gap aligns with the goal of data visualization: to aid human analysis and judgment with data. Visualization combines the inherent strengths of our visual system with the powerful graphical and computational abilities of the computer. A properly designed visualization allows our eyes to quickly discern patterns in data, which feeds our understanding of the underlying features and phenomena in our data. Visual inspection leads us to new insights about our data, helping us form hypotheses about where to focus next. Interaction allows us to further pursue these hypotheses either by showing other parts of the data, or by showing the same data from a different perspective. These features make data visualization an invaluable tool for exploration, analysis, and communication.

Effective data visualization is difficult, however. Most of us are familiar with basic graphs and charts. Given the prevalence of tools like Microsoft Excel, it would be difficult to find a colleague who has never spent time creating bar charts and pie charts. If you were to ask the same colleague to explain whether their bar chart is superior to a pie chart, however, no answer they give would be complete. The relative effectiveness of even the most basic charts is still a topic of debate. (Most research suggests that pie charts are the inferior choice, though some disagree, and recent work has begun investigation just how people read them.) Charts common in more technical contexts are no exception: histograms, box-plots, and scatterplots are frequently the topic of ongoing studies.

Given its ongoing study, one might wonder exactly when data visualization is useful in cyber security. Examples are readily available. When analyzing system-

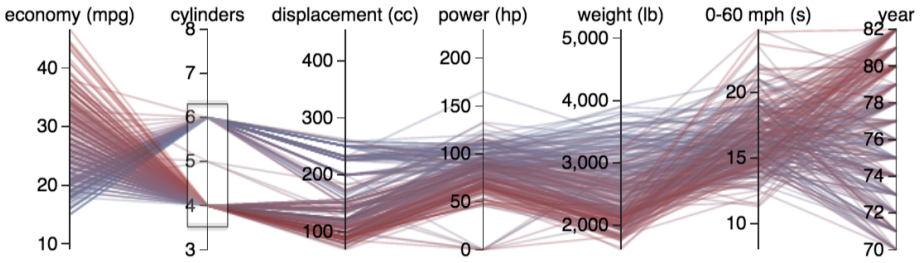

Figure 1: A parallel-coordinates plot [1]. Each axis represents a dimension (column) of the data. Each line represents a row (a particular car, in this case).

level logs, for instance, two of the most commonly used tools are command-line utilities and Excel. Command-line utilties are used to access, manipulate, and filter the logs, while Excel is used for inspection, analysis, and visualization. SIWs use these visualizations not only to aid in their analysis, but also to communicate their results to other security teams and stakeholders.

Examples of techniques developed through data visualization research are also plentiful. For instance, the data visualization capabilities of tools like Excel rapidly become difficult when the number of columns is large. If an analyst needed to look for relationships across 20 columns, for instance, using Excel they would need to manually create multiple charts comparing pairs of columns. Data visualization research offers several scalable alternatives. One is the parallel-coordinates plot (see Figure 1), which shows multiple dimensions side-by-side, and affords several interaction techniques to allow users to arrange columns to look for hidden correlations and outliers.

Cyber security is much more than log analysis, however. SIWs handle everything from threat and vulnerability management to forensics, traffic analysis, and more. As we cover these topics in the remainder of the chapter, bear in mind that our focus is not on covering the entire space, but rather on a sample of applicable data visualization techniques for each area.

2 Threat Identification, Analysis, & Mitigation

Threat analysis involves identifying and categorizing actions (both purposeful and accidental) that could interrupt day-to-day operations. Threat analysis uses many data sources, for instance the risk of host or section of the network on its

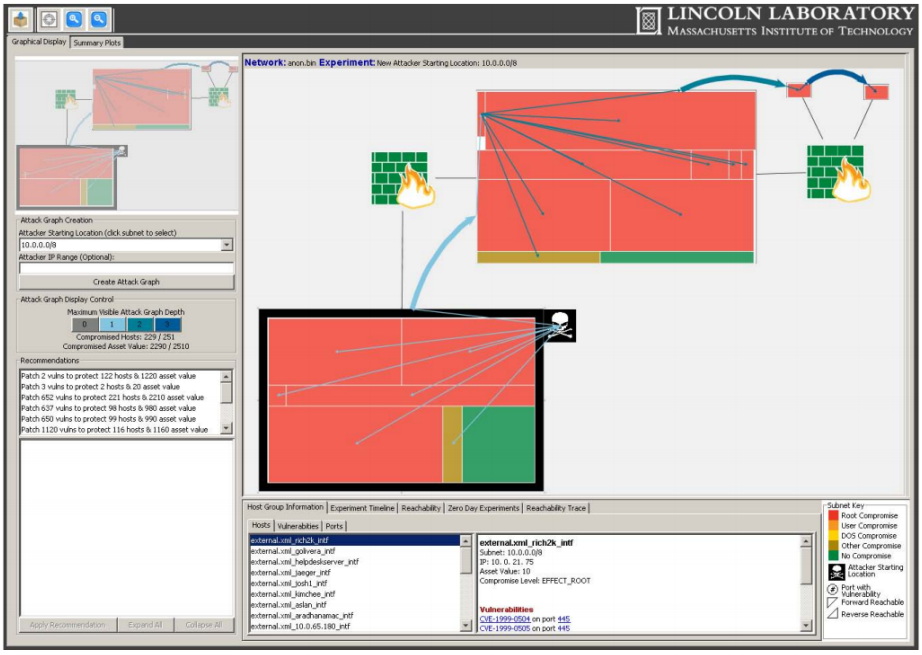

Figure 2: The NAVIGATOR system from Chu *et al.* [4]. Attack graph reachability information is combined with vulnerability information through node-link diagrams with treemap diagrams as the nodes.

known vulnerabilities, as well as data on the possible paths an attacker might take through a network or into a system. SIWs spend a significant amount of time on threat analysis in their organizations because it is the primary means of prioritizing ongoing maintenance and response procedures in the event of a breach.

ATTACK GRAPHS Diagrams can help illustrate the possible ways an attacker can traverse through a network towards a high value target. Commonly called attack graphs, these diagrams enable SIWs to computationally quantify threat analysis. These graphs can be computationally constructed from several pieces of information available in organizations, such as the network structure, the systems running on the network, and their associated vulnerabilities. Attack graphs allow SIWs to identify necessary changes in their networks and systems, and to test whether their changes were actually successful in stopping potential attacks.

From a data analysis perspective, the challenge with attack graphs is that they quickly become large and multi-faceted. There are usually multiple possible paths (edges) between systems (nodes) in a network. Each of these edges and nodes

may have several attributes, leading to a large multi-dimensional graph. Visual representations of large graphs often lead to a "hairball", where the large number of node and edge-crossings make it difficult to see the underlying structure of the graph.

Challenges in graph visualization are a perennial topic in data visualization research. In security visualization, several recent approaches have focused on showcasing how careful encoding and interaction design can improve cyber analysis with attack graphs.

Generally, attack graphs are represented as node-link diagrams and use different colors or shapes for the nodes and links to represent risk, vulnerability-type, and other available variables [27, 42, 28, 13]. Since node-link diagrams are often one of the best visual representations for following links [8], analysts are able to explore how attackers might gain access to critical machines through existing vulnerabilities. Node-link diagrams have several well-documented limitations, however [8], so researchers have proposed novel visual metaphors for attack graphs, including pixel-based matrix views and treemaps [26, 4] (see Figure 3).

Given the complexity of visualizing graphs, each attack graph technique has relative strengths and limitations. SIWs may benefit from having several of these techniques at their disposal.

Configuration There are several research efforts that attempt to make the configuration of security assets, particularly firewalls and routers, easier for SIWs.

Firewall configuration remains an open and difficult problem for security analysts. Rules can conflict in a number of ways, and firewall configuration remains computationally intractable. Therefore, effective firewall configuration in large networks may benefit through a hybrid approach that combines configuration algorithms with human-centered visualization techniques. Recent developments in this area have visualized different types of rule conflicts [24], and focused on accurately representing the dimensions of firewall configuration [22].

It is also important for analysts to understand router configuration. This problem differs from firewalls in that it requires the analysis of large traffic data, reachability data, routing data, and more. Basic visualizations such as scatterplots and stacked graphs can help SIWs identify patterns in DNS traffic [32]. Given the diverse data types, coordinated multiple view visualizations have been shown useful in representing these data sources together, allowing analysts to find important connections in the data [38, 33].

Navigation Orienting threat analysis becomes a challenge as organizations

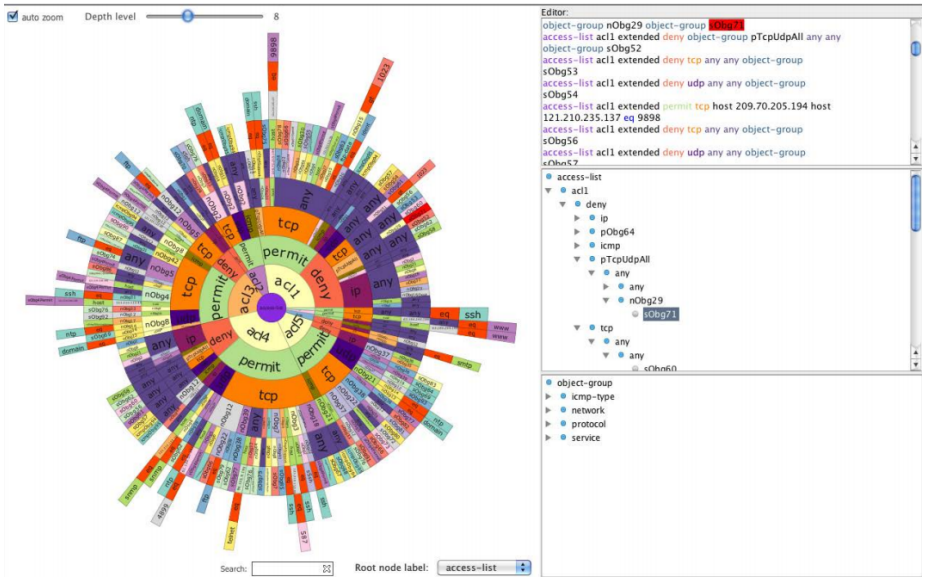

Figure 3: Firewall rule-set visualization from Mansmann *et al.* [22]. Hierarchically-evaluated rules are arranged in a sunburst diagram. Rule predicates towards the inside of the sunburst are processed first, while those on the outside are processed at the end (if reached at all).

grow and add new data sources. This growth increases the size and complexity of data that security analysts must navigate in their day-to-day activities. Recognizing this trend, several researchers have focused on visualization techniques that aid security analysts in navigating complex heterogenous data.

The mantra "Overview, zoom and filter, details on demand" is one of the most widely known and followed design guidelines in interactive data visualization [34]. An overview of data is usually intuitive. For a company focused on sales, it may be sales and related metrics over the past year, perhaps shown weekly or a month at a time. Overviews like this provide a starting point for an analyst at the company to zoom and filter into daily or even individual transactions of interest.

In cyber security, however, choosing a good overview is more difficult. Given the large and varied sources of data in an organization, an overview that serves as a starting point for all the types of analysis done in a security setting is not possible. This complexity has led to three commonly chosen overviews: the node-link diagram, the geospatial map, and the dashboard [15].

Many overviews start with the *topology* of the network. Network topologies are naturally represented as a node-link diagram, with individual IPs as nodes and connections between them as edges. Such data is readily inferred from a variety of widely available network data, including netflow, pcap, and firewall logs, for example.

Node-link diagrams, despite their prevalence in cyber security, suffer from many well-documented limitations [8]. Seeing a need for better approaches to overviews in security settings, Fowler and colleagues created IMap – a map-based overview of a large network [7].

The IMap algorithm begins with a node-link representation of a network, and leverages this structure to build a geographic map, like the familiar maps we use for directions. The resulting maps resemble continents with country boundaries, indicating different portions of the network. This transformation eliminates several problematic features of node-link diagrams, like hairballs of edge-crossings and difficult-to-decipher positions of nodes.

Beyond mitigating the tradeoffs of representations commonly used for overviews, map-inspired representations bring several immediate benefits to analysts. For one, they capitalize on these results and other research demonstrating that two-dimensional spatializations of data aid navigation [39]. Another way to think of IMap is as an even higher-level overview than the node-link diagrams typically used. This is true by definition, as the authors demonstrate that individual nodes

and links can be shown on demand from inside the IMap. Further benefits have yet to be investigated. However, decades of research in cartography have pointed to spatial maps as a useful tool for helping analysts find "common ground", and orienting themselves when facing new data and scenarios [21].

Now having discussed overviews, consider a situation in which an analyst has used an overview to identify activity of interest in their network. Navigation is still a challenge in this situation, because the analyst must identify the context in which the activity occurred to determine what actions are appropriate to take. In this case, context might include information like what subnet the activity was on, which user was logged in, and what machines are in the "vicinity" of the activity (in a connectivity sense, not necesarily meaning that the machines are physically close).

Analysts must reason about the context of suspicious activity when deciding how to respond. Few tools explicitly support this type of reasoning, however. An analyst might use tools to trace the route of one machine to another to discover other IPs of interest, yet this information must also be combined with context about the IPs along the route themselves.

Recognizing this need, Gray *et al.* have contributed the notion of Contextual Network Navigation [11]. In their work, the primary activity of interest is placed at the center of a network (*i.e.* egocentric), while other machines of interest are placed around it. This context can extend well beyond an organizational network and into the internet. Context outside of a network can be useful in helping analysts reason about the origin of attacks, for instance, providing information that can be used to strengthen the borders of the network.

3 Vulnerability Management

In order to assess the security posture of the servers and workstations on their network, security analysts and systems administrators use vulnerability assessment tools to scan their network and deployed code for potential exploits. In an organization with many services and systems, the number of open vulnerabilities is typically large and constantly changing.

It is unreasonable to expect all vulnerabilities to be patched – employees introduce new software to their machines weekly, and even new versions of trusted software can introduce vunerablities. SIWs must prioritize where they spend their

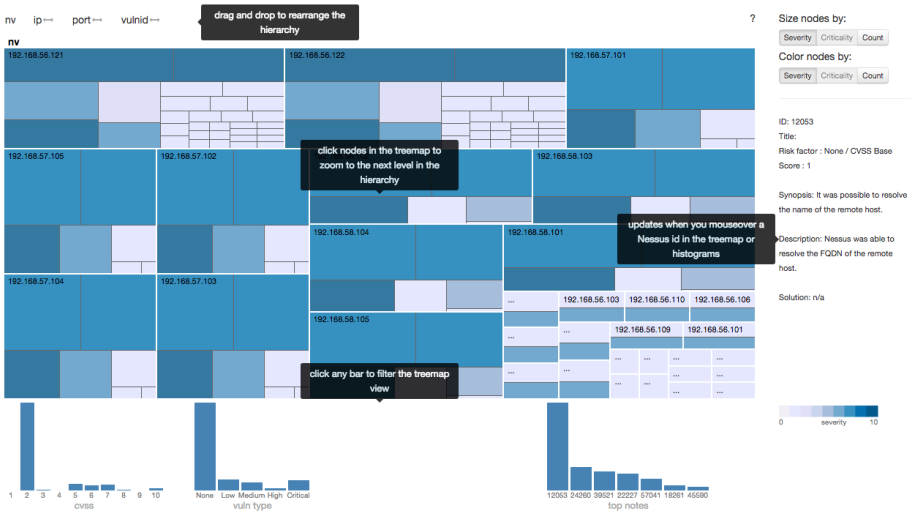

Figure 4: NV system from Harrison *et al.* [12]. A vulnerability scan of multiple machines at once is shown in a treemap diagram. Users can re-arrange the hierarchy to focus on ports, IPs, or individual vulnerabilities.

time. Relying on lists of vulnerabilities and SIW's mental models of a network to determine what should be prioritized can lead to troubling results, however. Industry reports show that unpatched systems are to blame for some of the largest and most harmful breaches in recent years [37].

NETWORK VULNERABILITIES Nessus and similar tools can produce an overwhelming amount of data for large networks. Network vulnerability scanners probe machines to determine which network ports are open, what services are running on the ports, and, most importantly, what versions of those services are running. Identifying the services and the versions enables these tools to match them with known vulnerabilities.

Scan analysis tools usually present the data in tables, sometimes with color coding to attempt to provide an overview of each vulnerability's severity. But scan data can be very large. With little support for comparing individual or logical groupings of machines, it can be difficult for SIWs to build a mental picture of what the overall vulnerability status is in the network. Further, it can be difficult to determine how the vulnerability status of a network has changed between scans at different points in time.

Nv uses treemaps and linked histograms to allow security analysts and systems

administrators to discover, analyze, and manage vulnerabilities on their networks [12] (see Figure 4). In addition to visualizing single Nessus scans, nv supports the analysis of sequential scans by showing which vulnerabilities have been fixed, remain open, or are newly discovered.

Software Vulnerabilities Vulnerabilities are not limited to commercial software. Organizations may run a number of scripts and services in their network for business operations. In fact, one of the main concerns of SIWs is discovering the "known unknowns" of the threats that exist in their organization's infrastructure [37].

Organizations must detect and manage vulnerabilities not only in the software they deploy, but also in the software they make. Similar to network vulnerability scanning, there are many tools for detecting and logging vulnerabilities in a given source code.

Recognizing this need, Goodall *et al.* developed a visual analysis tool to help understand and correct vulnerabilities in code bases [10]. Their system includes multiple views of not only the vulnerabilites detected by logging tools, but also the code itself. Exploiting the simplicity of the problem allows a unique self-contained system in which analysts can prioritize and fix vulnerabilities in a single tool.

Vulnerability management and remediation make up a large portion of how SIWs spend their time. There are few data-oriented tools and workflows specifically targeted at vulnerability management, however– this is an opportunity for future work.

4 Forensics

Visualization can benefit forensic analysis after a network or system is determined to be compromised. Forensic analysis requires in-depth data from one or several machines and devices. While these data sources can often be visualized with traffic and activity monitoring tools, the views in these tools are often of limited use in forensics, where the SIW needs to build a story of how an attack occurred. A distinguishing feature of forensics toolsis the lower level of detail they present to analysts. Research focusing on forensic analysis has dealt with either network data, individual device data, or behavioral data (*e.g.* email) in a post-intrusion setting.

Forensics tools focus on smaller subsets of the network, favoring features that show more detail and allow analysts to pivot between machines or subnetworks.

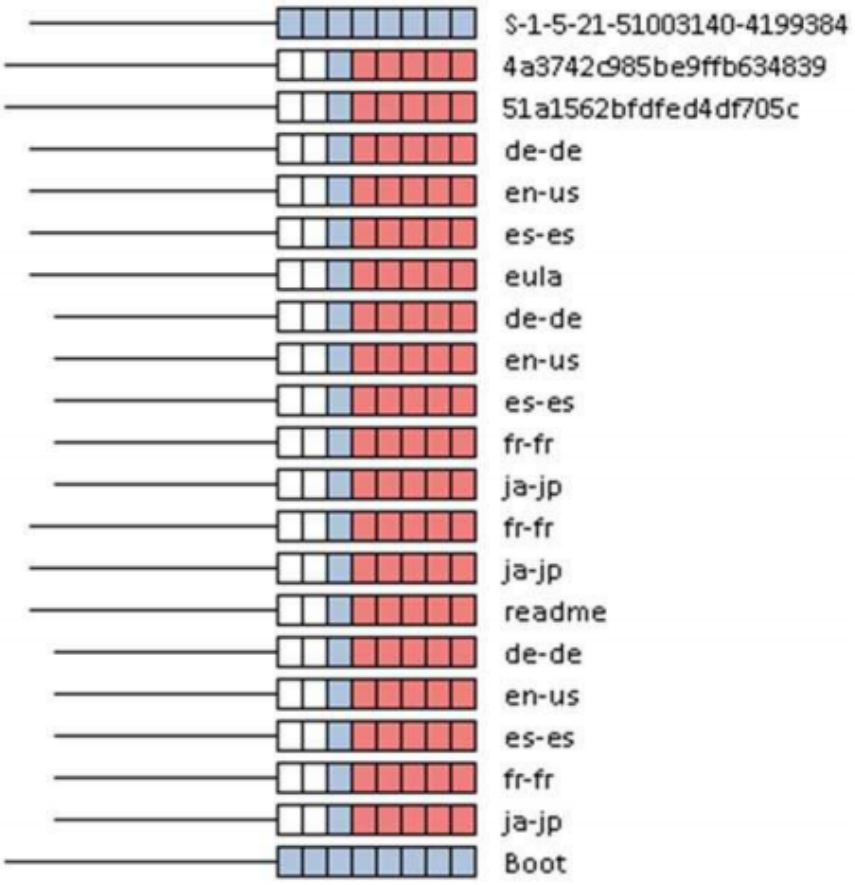

Figure 5: Change-Link system from Leschke *et al.* [19]. Malicious processes and patches often change directory structures temporarily, deleting evidence of their work as they finish. Change-Link shows how directories have been changed over time.

For example, instead of relying solely on text-based command-line utilities like grep, detailed search results can be visualized to assist analysts in exploring and filtering historical network traffic and events [18] . Several approaches have combined visualization with machine-learning to create and refine classification models for historic data, which can then be used to aid future forensics efforts [25, 43].

Visualization has also supported device-level forensics, particularly in digital string search, a core component of forensic analysis. In particular, traditional search algorithms have been used in conjunction with visual techniques to provide an overview of the search results and to display file and string interactively [17]. Another approach is to visualize changes to directory trees, as many attacks can be identified by how they modify files on a system [19].

Other applications have explored how visualization can benefit the behavioral forensics, such as attacks that involve email and chat. Email flow and group analysis results have been visualized to provide insight into both a users' typical usage patterns history [20] and to uncover malicious spam campaigns [40].

5 Traffic

Many data visualization tools in cyber security are designed to facilitate network monitoring and intrusion detection, or *situation awareness*. Situation awareness refers to knowing what is happening in a specific, rapidly changing environment.

Situation awareness can be broken down into three stages: perception, comprehension, and projection [6]. Perception involves monitoring the status of a network and systems within a network, while comprehension involves synthesizing the status of the individual elements in the first stage to recognize patterns that should be investigated further. In contrast, projection involves extrapolating the current state of the network towards the immediate future. In security visualization, most efforts and tools target perception and comprehension.

There are several stages of network analysis and defense [5]. These include monitoring and initial detection, as well as the network forensics and reporting and decision making activities that follow. Other studies have examined how analysts make decisions and how visualization might benefit that process [31], as well as how high-level security goals relate to low-level data [14].

Another focus area in situation awareness research is to examine analysts'

existing toolkits to identify the strengths and limitations of current tools and workflows. These studies lay the groundwork for new visualization techniques and systems. For instance, although command-line search tools such as grep remain a staple in the analyst's toolkit, yet few visualization systems incorporate even basic search functionality [3], potentially damanging their adoption. Similarly, analysts often need to relate external security information to activity on their network (*e.g.* from websites, mailing lists, and other online sources). Recent data visualization systems have begun to process external information and visualize how it relates to internal network operations [30].

Visualization designs have taken many forms in support of Situation Awareness. Given the time-focused nature of network security data, some visualization techniques emphasize time and support interactions that allow SIWs to move forward and backward through the temporal dimensions of their data. One such approach is event plots, which visually encodes events by category using visual marks to represent events [29]. Challenges for event plots include defining what an event is (*e.g.* packets are low-level events, whereas emails are highlevel events) and defining how to represent an event (*e.g.* color, shape). While clutter can become an issue, event plots becoming more widely used because they can provide both an overview and a detailed view of a large number of machines.

6 Emerging Themes

Data visualization has been applied to many areas of security analysis. Security analysis is changing, however. Two emerging themes are a push to move analysis towards intelligent systems, and a realization that data analysis benefits not only the individual SIW, but their team, their organization, and their community.

HUMAN-IN-THE-LOOP MACHINE LEARNING Machine learning has had considerable success in some areas of cyber security, but limited success in others. SPAM filters are a machine learning success story. With numerous examples of spam emails, machine learning is used to accurately vet incoming mail for spam-like features (most of the time, at least). This reduces the number of phishing emails that make it to end-users, considerably improving the security of the organization. Yet as we have seen, SIWs focus on much more than email.

Machine learning has been repeatedly applied towards helping SIWs identify anomalous traffic and connections in their networks. These efforts have not seen widespread adoption, however, for several reasons: Attacks are rare, meaning the

number of false positives a model produces will be much larger than the actual number of attacks [2]. (In spam the situation is different, given the large number of malicious emails.) This false alarm problem is multiplied when machine-learning models are built using different data sources or sub-networks. Smaller models like this are often necessary for building a model of what normal activity looks like.

High-frequency data such as traffic and connections in a network can lead to an overwhelming number of false positives that SIWs must investigate. In these cases, the analysts' ability to defend the network is significantly reduced [14, 31].

A deeper problem is the understandability of the models produced by machine learning. While some learning algorithms produce human-readable output, a sequence of questions about features in the data for example, other algorithms produce models that are difficult for humans to internalize. This is a general problem in machine learning, but because security events must be vetted by SIWs, the gap between humans and models in machine learning remains a longstanding challenge [2, 16, 35].

Given the potential impact, researchers have been actively working to close this gap. One approach is to add constraints to the learned models that align with SIW's ability to understand them [9]. Another emerging approach is to conduct assessments of the tools SIWs would need to better manage multiple models [41]. These approaches hold promise to meld the operational needs of SIWs with the structure, capabilities, and limitations of machine learning.

Exploration versus exposition Security has become more data-driven in recent years. Data manipulation tools have become more usable, analysis techniques have become more scalable, and visualization techniques have become more aligned with the strengths and limitations of our perceptual and cognitive abilities. But data analysis and exploration is only part of the story.

Beyond exploration, the communication and exposition of data has become a central topic in security. Part of this is driven by data-laden industry reports. The Verizon Data-Breach Investigations Report [37], for example, collects breach reports from multiple organizations and identifies emerging trends and gaps in the past year. Reports such as these are invaluable as organizations plan and pull-together resources for future years.

Given the expansion beyond exploratory data visualization to expository data visualization, new challenges have arrived. Best practices such as only using the absolute best perceived visual stimuli in data visualization (such as position or length) make sense in exploratory contexts where it is assumed that you have the

full attention of the viewer. In exposition, however, accuracy in visual stimuli may need be sacrificed for stimuli that are engaging.

Unfortunately, many practitioners, companies, and academics remain unaware of best practices in data visualization. This leads to a number of flashy (yet ineffective) analysis tools and attractive (yet misleading) industry reports. More study is necessary to navigate this space. As research moves from focusing on techniques to focusing on the SIW [36, 23], security visualization will become an even more integral part of day-to-day security operations and communication.

References

[1] Parallel coordinates. https://syntagmatic.github.io/parallel-coordinates/. Accessed: 2016-06-01.

[2] S. Axelsson. The base-rate fallacy and the difficulty of intrusion detection. *ACM Transactions on Information and System Security (TISSEC)*, 2000.

[3] S. Bratus, A. Hansen, F. Pellacini, and A. Shubina. Backhoe, a packet trace and log browser. In *Visualization for Computer Security*. 2008.

[4] M. Chu, K. Ingols, R. Lippmann, S. Webster, and S. Boyer. Visualizing attack graphs, reachability, and trust relationships with navigator. In *Proceedings of the Seventh International Symposium on Visualization for Cyber Security*, pages 22–33. ACM, 2010.

[5] A. D'Amico and M. Kocka. Information assurance visualizations for specific stages of situational awareness and intended uses: lessons learned. In *Visualization for Computer Security, 2005.(VizSEC 05). IEEE Workshop on*, 2005.

[6] M. R. Endsley. Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 1995.

[7] J. J. Fowler, T. Johnson, P. Simonetto, M. Schneider, C. Acedo, S. Kobourov, and L. Lazos. Imap: Visualizing network activity over internet maps. In *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, 2014.

[8] M. Ghoniem, J.-D. Fekete, and P. Castagliola. A comparison of the readability of graphs using node-link and matrix-based representations. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*, 2004.

[9] M. Gleicher. Explainers: Expert explorations with crafted projections. *Visualization and Computer Graphics, IEEE Transactions on*, 2013.

[10] J. R. Goodall, H. Radwan, and L. Halseth. Visual analysis of code security. In *Proceedings of the seventh international symposium on visualization for cyber security*, 2010.

[11] C. C. Gray, P. D. Ritsos, and J. C. Roberts. Contextual network navigation to provide situational awareness for network administrators. In *Visualization for Cyber Security (VizSec), 2015 IEEE Symposium on*, 2015.

[12] L. Harrison, R. Spahn, M. Iannacone, E. Downing, and J. R. Goodall. Nv: Nessus vulnerability visualization for the web. In *Proceedings of the ninth international symposium on visualization for cyber security*, 2012.

[13] J. Homer, A. Varikuti, X. Ou, and M. A. McQueen. Improving attack graph visualization through data reduction and attack grouping. In *Visualization for Computer Security*, pages 68–79. Springer, 2008.

[14] C. Horn and A. D'Amico. Visual analysis of goal-directed network defense decisions. In *Proceedings of the 8th international symposium on visualization for cyber security*, 2011.

[15] J. Jacobs and B. Rudis. *Data-driven Security: Analysis, Visualization and Dashboards*. John Wiley & Sons, 2014.

[16] P. Jaferian, D. Botta, F. Raja, K. Hawkey, and K. Beznosov. Guidelines for designing it security management tools. In *Proceedings of the 2nd ACM Symposium on Computer Human interaction For Management of information Technology*, 2008.

[17] T. Jankun-Kelly, D. Wilson, A. S. Stamps, J. Franck, J. Carver, J. E. Swan, et al. A visual analytic framework for exploring relationships in textual contents of digital forensics evidence. In *Visualization for Cyber Security, 2009. VizSec 2009. 6th International Workshop on*, 2009.

[18] K. Lakkaraju, R. Bearavolu, A. Slagell, W. Yurcik, and S. North. Closing-the-loop in nvisionip: Integrating discovery and search in security visualizations. In *Visualization for Computer Security, 2005.(VizSEC 05). IEEE Workshop on*, 2005.

[19] T. R. Leschke and A. T. Sherman. Change-link: a digital forensic tool for visualizing changes to directory trees. In *Proceedings of the ninth international symposium on visualization for cyber security*, 2012.

[20] W.-J. Li, S. Hershkop, and S. J. Stolfo. Email archive analysis through graphical visualization. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, 2004.

[21] A. M. MacEachren. *How maps work: representation, visualization, and design*. Guilford Press, 1995.

[22] F. Mansmann, T. Göbel, and W. Cheswick. Visual analysis of complex firewall configurations. In *Proceedings of the ninth international symposium on visualization for cyber security*, 2012.

[23] S. McKenna, D. Staheli, and M. Meyer. Unlocking user-centered design methods for building cyber security visualizations. In *Visualization for Cyber Security (VizSec), 2015 IEEE Symposium on*, 2015.

[24] S. P. Morrissey and G. Grinstein. Visualizing firewall configurations using created voids. In *Visualization for Cyber Security, 2009. VizSec 2009. 6th International Workshop on*, 2009.

[25] C. Muelder, K.-L. Ma, and T. Bartoletti. A visualization methodology for characterization of network scans. In *Visualization for Computer Security, 2005.(VizSEC 05). IEEE Workshop on*, 2005.

[26] S. Noel, M. Jacobs, P. Kalapa, and S. Jajodia. Multiple coordinated views for network attack graphs. In *Visualization for Computer Security, 2005.(VizSEC 05). IEEE Workshop on*, pages 99–106. IEEE, 2005.

[27] S. Noel and S. Jajodia. Managing attack graph complexity through visual hierarchical aggregation. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, pages 109–118. ACM, 2004.

[28] S. O'Hare, S. Noel, and K. Prole. A graph-theoretic visualization approach to network risk analysis. In *Visualization for computer security*, pages 60–67. Springer, 2008.

[29] D. Phan, J. Gerth, M. Lee, A. Paepcke, and T. Winograd. Visual analysis of network flow data with timelines and event plots. In *VizSEC 2007*. 2008.

[30] W. A. Pike, C. Scherrer, and S. Zabriskie. Putting security in context: Visual correlation of network activity with real-world information. In *VizSEC 2007*. 2008.

[31] J. Rasmussen, K. Ehrlich, S. Ross, S. Kirk, D. Gruen, and J. Patterson. Nimble cybersecurity incident management through visualization and defensible recommendations. In *Proceedings of the Seventh International Symposium on Visualization for Cyber Security*, 2010.

[32] P. Ren, J. Kristoff, and B. Gooch. Visualizing dns traffic. In *Proceedings of the 3rd international workshop on Visualization for computer security*, 2006.

[33] J. Shearer, K.-L. Ma, and T. Kohlenberg. Bgpeep: An ip-space centered view for internet routing data. In *Visualization for Computer Security*. 2008.

[34] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, 1996.

[35] R. Sommer and V. Paxson. Outside the closed world: On using machine learning for network intrusion detection. In *Security and Privacy (SP), 2010 IEEE Symposium on*, 2010.

[36] D. Staheli, T. Yu, R. J. Crouser, S. Damodaran, K. Nam, D. O'Gwynn, S. McKenna, and L. Harrison. Visualization evaluation for cyber security: Trends and future directions. In *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, 2014.

[37] V. R. Team et al. Verizon data breach investigations report, 2016.

[38] S. T. Teoh, S. Ranjan, A. Nucci, and C.-N. Chuah. Bgp eye: a new visualization tool for real-time detection and analysis of bgp anomalies. In *Proceedings of the 3rd international workshop on Visualization for computer security*, 2006.

[39] M. Tory, D. W. Sprague, F. Wu, W. Y. So, and T. Munzner. Spatialization design: Comparing points and landscapes. *Visualization and Computer Graphics, IEEE Transactions on*, 2007.

[40] O. Tsigkas, O. Thonnard, and D. Tzovaras. Visual spam campaigns analysis using abstract graphs representation. In *Proceedings of the ninth international symposium on visualization for cyber security*, 2012.

[41] S. Walton, E. Maguire, and M. Chen. A visual analytics loop for supporting model development. In *Visualization for Cyber Security (VizSec), 2015 IEEE Symposium on*, 2015.

[42] L. Williams, R. Lippmann, and K. Ingols. *GARNET: A graphical attack graph and reachability network evaluation tool*. 2008.

[43] C. Wright, F. Monrose, and G. M. Masson. Hmm profiles for network traffic classification. In *Proceedings of the 2004 ACM workshop on Visualization and data mining for computer security*, 2004.