Eurographics Conference on Visualization (EuroVis) 2016
K-L. Ma, G. Santucci, and J. J. van Wijk
(Guest Editors)

Volume 35 (2016), Number 3

BubbleNet: A Cyber Security Dashboard for Visualizing Patterns

S. McKenna[1,2], D. Staheli[2], C. Fulcher[2], and M. Meyer[1]

[1]University of Utah
[2]MIT Lincoln Laboratory

Abstract

The field of cyber security is faced with ever-expanding amounts of data and a constant barrage of cyber attacks. Within this space, we have designed BubbleNet as a cyber security dashboard to help network analysts identify and summarize patterns within the data. This design study faced a range of interesting constraints from limited time with various expert users and working with users beyond the network analyst, such as network managers. To overcome these constraints, the design study employed a user-centered design process and a variety of methods to incorporate user feedback throughout the design of BubbleNet. This approach resulted in a successfully evaluated dashboard with users and further deployments of these ideas in both research and operational environments. By explaining these methods and the process, it can benefit future visualization designers to help overcome similar challenges in cyber security or alternative domains.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—User-centered design

1. Introduction

Over the past ten years, roughly two *billion* pieces of digitized personal information have been lost or stolen, largely by hackers [Kas15]. Several note-worthy breaches include: Sony Pictures discovered that over one-hundred terabytes of data ranging from films to employee information to sensitive business documents were copied off of their networks; personal information such as names, addresses, phone numbers, and emails were found by hackers with administrative access to the US's largest bank, JP Morgan Chase; T-Mobile customers have had sensitive personal information leaked from a breach within the Experian credit agency, everything from names to social security and passport numbers.

Hacks like these are becoming increasingly prevalent and sophisticated, making the maintenance of a safe and secure computer network challenging, yet critical. Maintaining security on these computer networks is extremely challenging, particularly due to the scale of the data as well as the constantly evolving nature of cyber security attacks [EFC*10, BEK14]. Often, these attacks require a human interpretation in order to uncover, stop, and recover from these attacks [DW08]. Network analysts struggle with a very data-intensive task where it is easy to make mistakes, errors, and miscalculations [EFC*10]. Visualization is one way for analysts to both explore and present this large data space, but analysts have been known to be hesitant about trusting visualizations for their own workflows [FNER09].

In this paper we describe a design study focusing on the domain of cyber security, where we worked with two dozen different cyber security experts over the span of two years with the goal of improving how analysts discover and present interesting anomalies and patterns within computer network data. To the best of our knowledge, this is the first end-to-end design study within this domain. Conducting the design study brought about an interesting set of design constraints: limited access to the analysts and data, multiple types of end-users, and deployment limitations. Some of these challenges go against guidelines for conducting design studies, such as arguments for an up-front winnowing of users and collection of data [SMM12]. Addressing these issues, however, allowed us to validate a number of other guidelines for incorporating user-centered design methods into a cyber security project [MSM15], as well as for making use of a variety of discourse channels [WBD14].

The primary contribution of this design study is the design, evaluation, and deployment of an interactive dashboard, BubbleNet, for visualizing patterns in cyber security data. BubbleNet is designed to not only support the discovery of patterns, but to facilitate presentation of these patterns to various stakeholders. We discuss a problem characterization for this domain, along with a data and task abstraction. A secondary contribution of this work is a detailed discussion of the design process, including use of several different user-centered design methods [MSM15], as well as an application of the channels of discourse strategy [WBD14].

In the first part of this paper we compare against related work in

Section 2 and describe the data and task abstraction in Section 3. Next, Section 4 discusses a methodical design process for the unique design constraints we encountered. The BubbleNet dashboard is described in Section 5, which Section 6 discusses evaluation both as a usability study and deploying to real users. Lastly, Section 7 highlights implications from what we have learned and how they apply to both the cyber security and visualization communities.

2. Related Work

The tasks of discovery and presentation are open challenges in terms of visualization for cyber security. Many visualization tools and techniques are designed to fit the data, not the users [SYC*14]. Furthermore, the visualization and cyber security research is largely evaluated with use-cases involving toy datasets and researchers not practitioners in the field [SYC*14]. In addition, very few tools have considered how to present cyber information to stakeholders with less technical experience and knowledge, such as IT personnel or network managers. Large organizations often have analysts working together in teams and with a variety of other individuals, such as their managers, in order to convey priorities and matters of importance to those in leadership roles who make decisions [MSM15, EFW*10].

Numerous cyber security researchers have adapted existing visualizations for data in this domain, but very little of this work has tested the usability or utility for network analysts. Different researchers have plotted cyber security data on bar and scatter plots [HHH13, Wat05, LYL04]. Other researchers have explored using a heatmap or matrix to encode various attributes and hierarchies within the data [Wat05, KOK05, LHS14, KRA*05]. Parallel coordinates have also been utilized by several researchers to visualize multiple dimensions of data [ECS05, YYT*04, CLK09]. Goodall and Sowul went beyond a single parallel coordinates view with other details-on-demand visualizations like charts and maps into a simple dashboard [GS09]. There is potential to combine and link multiple visualizations together into a dashboard that is then evaluated against users.

Visualization research has sought out novel visual representations tailored to cyber security data. Network graph layouts have been adapted and focused within this domain [FAL*06, TPG*09, PRS*13]. Map-like visualizations of the entire internet seek to preserve spatial location of similar types of computers across multiple datasets [FJS*14]. Aggregated sliding slices of time is discussed by Fischer and Keim in order to support the workflow of network analysts dealing with large quantities of data [FK14]. While these techniques could be useful, most of them have not been evaluated with respect to their usability or effectiveness for network analysts with real data.

A number of cyber security researchers have studied the usability and effectiveness of their tools, but there is no common evaluation framework to utilize [SYC*14]. Researchers have developed custom surveys [MP08, FMN05, BDF*08, KRA*05], which make comparison difficult and may not account for response bias [Bro96]. Leschke and Nicholas evaluated a tool with a standardized usability survey [LN13] and others have performed formal user studies [RER*10, ALK*11], but none discuss deployment. Landstorfer et al. designed a visualization in a user-centered design process but only garnered initial user feedback [LHS14]. Hao et al. worked with analysts to showcase utility of web-based visualization dashboards for network security but did not evaluate with users' own data [HHH13]. While visualization researchers have worked with users, we have found no end-to-end design study in this space, from abstraction to deployment.

3. Problem Characterization and Abstraction

While most domain research focuses solely on data analysis, the task of presentation is a vital one for network analysts, as information must often be conveyed to other people for decisions to be made [MSM15]. Often, this information to convey and decisions to be made surround a problem or an incident [DW08]. One analyst we spoke with summarized why presentation is challenging: *"pictures are great when going up to management because you have 60 seconds to make your case"* (A4). There are many kinds of cyber security incidents which can result in negative outcomes, such as information disclosure, theft, and denial of service [HL98].

Cyber security includes a variety of data types such as logs of computer functionality, but network security is a subset which focuses on multiple computer interactions with a base unit of a **network record**. A *network record* is metadata associated with the communication between two computers. This metadata can include a whole variety of information such as time, location, priority, category, and various other attributes, collected from the details of the data such as the timestamp and IP address. There are a variety of different network security datasets, such as raw packet capture, net flow, intrusion detection systems, and firewall logs. Each of these datasets corresponds roughly with network records, but the key differences are the attributes or various metadata associated with each.

The basic unit of network security analysis is a **pattern**, a collection of network records that represent some recurring or abnormal behavior which can be benign or malicious. One way to create patterns is to summarize or aggregate records in different ways such as those coming from a specific computer, general location, or subsets of time. Benign patterns represent typical, authorized network records like typical outgoing web traffic along port 80. However, patterns can be malicious, such as a network scan from a single external computer in order to find vulnerabilities or disrupt an organization's network. These malicious patterns can be a collection of many network records like a network scan or even a single one where a hacker exfiltrates a sensitive document.

Pattern recognition and finding anomalies is a very crucial aspect for data science and machine learning in particular. Several researchers have adopted machine learning techniques for cyber security [ALK*11] and also for finding anomalies in social media analysis [ZCW*14, CSL*16]. All of these authors discuss the rich and deep applications of machine learning for each of these domains. Due to the large scales of data in cyber security, these techniques can and often are utilized to find subsets of potentially interesting network records to visualize, but humans are often still required to analyze these results and are a critical component of this triage process [ALK*11].

Another way to formulate patterns is to consider different aggregations of network records, like time and location. Many cyber security visualizations have been developed for showing hierarchical time-varying aspects of the data [ECS05, FK14]. From working with users, we found that aggregation to a larger scale by *hours* and *days* is both useful and interesting. Network security datasets are commonly aggregated by IP address, and these can be visualized in many ways from IP grids to internet maps [GS09, LYL04, Wat05, KOK05, LHS14, KRA*05, ECS05, FJS*14]. Aggregation of computers can also occur by their location of an IP address, through databases like MaxMind GeoLite2 [Max15], used by other visualization tools, like EMBER [YLRB10]. We found that geolocation is the simplest and most intuitive way to present cyber information to different users. While not ideal, location can enable users to formulate patterns that correspond to geopolitical entities like *countries*. For visualizing anomalies, it is also useful to compute statistical information like averages.

For this design study, the task focus was on the discovery and presentation of cyber security patterns. Presentation of patterns requires simple and easily understood visualizations for consumption by users who are not domain experts. Discovery of patterns is an important part of network security analysis, encompassing tasks identified by previous researchers such as perception, detection, and monitoring [DW08]. Two different analysts equate discovering these patterns to finding a needle in a haystack, and the importance of aggregation is illustrated by this analyst's insight on our aggregation choices of hour, day, and country: *"we would have never have seen that [pattern] any other way, maybe if we even had [data] formatted a different way that pattern would have never emerged"* (A1). Finding patterns can be particularly challenging since cyber attackers are dynamic and constantly changing their methods. For both discovery and presentation, some important tasks include the ability to *identify* interesting patterns as well as *compare* patterns to find differences. For example, an interesting pattern could be activity at a certain hour of the day or a specific attribute between two different countries.

4. Design Process

This project focused on creating a dashboard for cyber security patterns. To present these patterns, there was a focus on users beyond the network analyst. As such, it was necessary to incorporate these other users, their needs, and workflows into the design process in order to create the final BubbleNet dashboard. This design process highlighted key insights into the connection, similarities, and differences of user-centered design and a design study, and these insights make this work unique compared to past user-centered design papers for cyber security. As a result, we reflected on this design process and modeled it in the form of Wood et al's discourse channels, which are *"complex relationship between producers and consumers of a visualization."* [WBD14] In this work, we utilized four distinct discourse channels: a software company, a research organization, university information security, and an operational organization. These different discourse channels interacted together and led to successful outcomes as a result.

We present an overview of our design process in Figure 1. Each row and color corresponds to a different discourse channel. Each of these channels have different users, data sources, and design methods that were employed. The primary outcomes of this process are the prototypes and tools, with screenshots shown above the timeline for each. Prototypes are linked via curved lines to evaluation methods, and the final BubbleNet dashboard in c) is linked to deployments in two different discourse channels.

The first portion of this design study was informed by a previous domain analysis: a qualitative coding of cognitive task analysis papers [MSM15]. In order to establish specific user needs, we performed a series of contextual semi-structured interviews at a research organization. As a result, four key user personas were identified for dashboard design [MSM15]. By evaluating project constraints, the project was further focused into two specific user personas: network analysts and managers.

After selecting this subset of users, user needs were adapted from a previous project [MMAM14] and prioritized against each of our user personas. Examples of these needs or *user requirements* include: scaling to real-world data on a single screen, preservation of data context, emphasizing temporal representations of patterns, designing visualizations for presenting to others, and keeping it both intuitive and easy to use. Next, two dozen different visualization encoding ideas were sketched and weighted against each need. As a result, each idea was scored by combining these priorities and weights, resulting in several key ideas with the most potential. We created the first prototype from these ideas, shown in Figure 1a). This prototype contained a treemap of network records, organized by city and country. We evaluated this prototype using Nielsen's usability heuristics and Gestalt principles. This method highlighted low-level changes, but we desired to evaluate the data abstraction and treemap encoding.

To perform this evaluation, we turned to the data sketches method [LD11]. Through existing tools and techniques, twenty different data sketches [MSM15] were shown to a collaborating network analyst to gather feedback on different encodings. This feedback discouraged us from using a treemap since it took significant time to present and explain these to an analyst. Furthermore, implementing the spatial treemap algorithm [WD08] uncovered tradeoffs between spatial location (topology) versus aspect ratio of each element (squarified). In other words, spatially relevant treemaps were more challenging to read and to compare size. For further detail on these sketches as both design alternatives and for an analysis tailored for the domain of cyber security, please see previous work on the data sketches design method [MSM15]. However, the feedback received on the data sketches validated our initial data abstraction of location-based aggregation since abstractions like network graphs are too complex for a simple summary view, whereas location-based views required little to no explanation.

Thus, we iteratively developed towards a location-based encoding which is simpler and more intuitive for a larger variety of users, shown in Figure 1b). A usability study was performed on this second prototype to evaluate its usability, and this resulted in the final BubbleNet dashboard in Figures 1c) and 2. While BubbleNet was deployed in a research environment, significant changes were necessary to create the final tool for deployment into an operational environment. These aspects of evaluation and deployment are discussed further in Section 6.

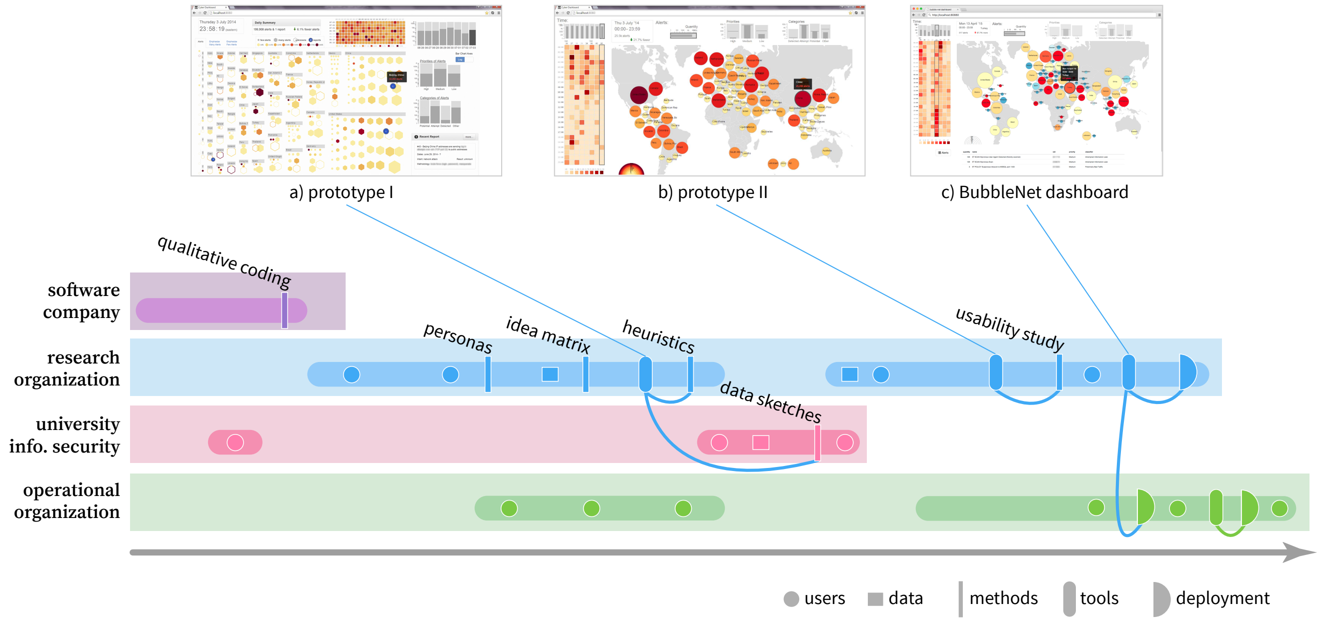


Figure 1: *An overview of our design process. Four distinct channels played a role in BubbleNet's design, the first was previous work, and the second and fourth were various users in two distinct settings, both research and operational. The third channel involved a network analyst from a university. Each channel involved different sets of users and data, but the final BubbleNet design in c) and deployments all occurred due to the interaction of outcomes and user feedback across all of these channels.*

5. BubbleNet Dashboard

We present the encodings and design justifications behind each view of the BubbleNet dashboard, shown in Figure 2. In BubbleNet, each view supports interactive selection of elements. This selection pivots the data in all other views on the fly to the given selection. This supports identifying interesting patterns and comparing them as well.

5.1. Location View

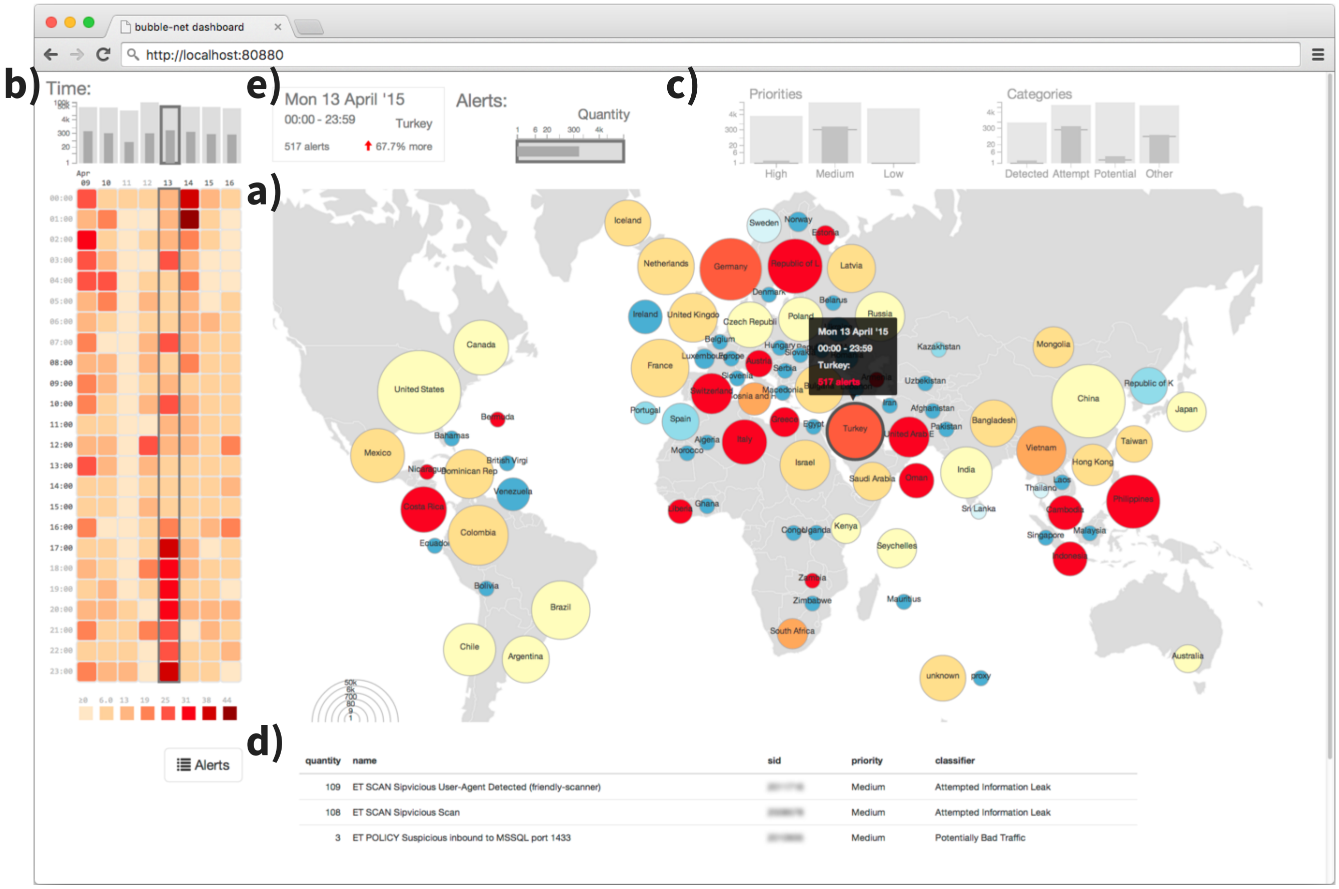
BubbleNet's primary view is a location-based map view shown in Figure 2a). This encoding is a Dorling-like cartogram [Bos15] which animates circles to preserve spatial location. The implementation here is a simplification of the Dorling cartogram algorithm [Dor11]. Each circle represents an aggregation of network records by country, and the Dorling-like cartogram is similar to a force-directed layout, initialized by the country centroids. Each circle is encoded in size by the quantity of records, and deviations from an average are encoded using color where red is more records than average and blue is less. Size is encoded on a log scale due to both the importance of visualizing a single record as well as the large range of record values, up to hundreds of thousands.

After gathering feedback on the initial treemap prototype, we learned that the details of the location (e.g. city) were less important and more uncertain to visualize in a single view. As discussed in the previous section, there are also caveats to utilizing a treemap algorithm since there are trade-offs between location and the squarified nature of the treemap. Furthermore, treemaps were not desired by us as designers due to aesthetic reasons of whitespace, since they are space-filling, unlike a map which has more whitespace. This is why the first prototype used hexagons instead of rectangles in the treemap in order to provide more whitespace between elements, but this was switched to circles since they are simpler and pack effectively on a map which utilizes whitespace more aesthetically to us.

Originally, the dashboard dual-encoded color and size to the number of records as in Figure 1b), but the usability study presented in Section 6 obtained requests from users to show change visually on the map. There were records which could not be geolocated via MaxMind [Max15], so they were placed on an empty portion of the map to save space. Interactions with various other views in the dashboard result in an animation of the force-directed layout algorithm, and these animated transitions did not appear to distract or annoy users but did captivate them. This animation enabled a more consistent map view for users, unlike the treemaps which resulted in more significant changes of size and location due to trade-offs of the underlying algorithms.

5.2. Temporal View

There are two views in Figure 2b) which encode time: a bar chart of network records per day with a common horizontal axis of days that aligns with a temporal heatmap beneath it where its vertical axis is by hour. The bar chart provides a quick overview of each day, and the heatmap provides details by the hour to support quick pattern discovery. It would be possible to derive similar encodings for different aggregates of time. The heatmap limits the number of days to a week in order to avoid data overload and reduce color



quantity	name	sid	priority	classifier
109	ET SCAN Sipvicious User-Agent Detected (friendly-scanner)	[redacted]	Medium	Attempted Information Leak
108	ET SCAN Sipvicious Scan	[redacted]	Medium	Attempted Information Leak
3	ET POLICY Suspicious inbound to MSSQL port 1433	[redacted]	Medium	Potentially Bad Traffic

Figure 2: *The BubbleNet dashboard labeled by its corresponding encodings: a) location map based on a Dorling cartogram, b) temporal chart and heatmap, c) attribute bullet bar charts, d) record details table, and e) selection overview.*

perception issues by keeping the heatmap squares larger. The bar chart and heatmap views are arranged along a common axis due to early user feedback and the heuristics evaluation, which resulted in moving, enlarging, and linking these two encodings to create an effective temporal pattern filter.

5.3. Attributes View

The BubbleNet dashboard also includes bar charts and bullet charts for different attributes of the data, e.g. the priority and category for each network record, shown in Figure 2c). Bullet charts are inspired by Stephen Few's bullet graphs for dashboards [Few10]. Bullet graphs encode a value, a qualitative ranking, an average, and a projection into a single element, but a *bullet chart* is simplified where an inner bullet represents a subset of the full bar. In other words, the entire world's value is represented as a lighter bar and the value of a selected country is the smaller, darker bullet inside it as in Figure 2c). Furthermore, the bullet chart similarly encodes the average for an individual country using a thin, dark line.

Bullet charts enable showing a subset of a larger value, i.e. a country's value with respect to the world's amount. Unlike bullet graphs, bullet charts show a quantitative subset, and this subset enables quick comparison through interaction. As with previous scales, we incorporated a log scale for these bar charts. Alternative encodings of the data were considered across all views, such as orders of magnitude markers [BDJ14], but these encodings required significant explanation and collided with encoding subsets. A log scale helps to visually show both extremely large and extremely small values at the cost of comparing values precisely, but interaction supports comparing precise values using text.

5.4. Records View

A details-on-demand table view in Figure 2d) provides a summary of the different records in any selection. This summary includes the quantity, user-friendly name, ID or type of record, and the detailed attribute information. These details enable analysts to understand what is happening in any selected aggregate of network records in the dashboard. As such, this table and dataset were created by request of all analysts during the usability study, presented in Section 6. Inclusion of network record details is critical to this discovery of patterns. In our evaluation, analysts told us that they were able to not only discover patterns using BubbleNet, but that they could envision using this dashboard to present what they found.

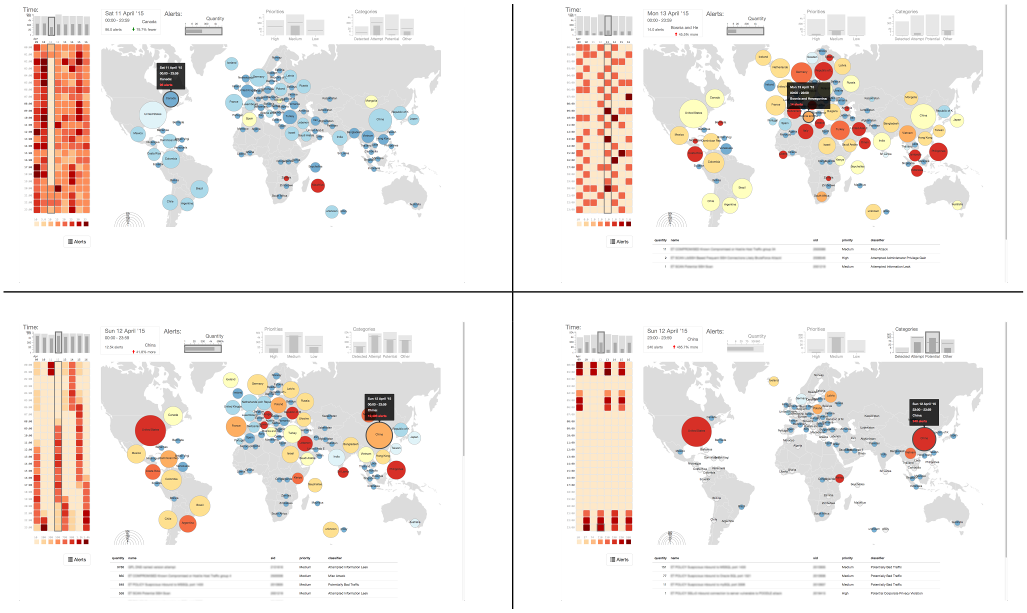


Figure 3: *Most elements of the BubbleNet dashboard are interactive and update all other views accordingly. For example, selecting four different countries shows significantly different patterns in the hourly heatmap.*

5.5. Selection and Interaction

Interaction is a crucial component of most elements on the BubbleNet dashboard. Most interactions involve a selection that specifies some pattern, which updates the selection window in Figure 2e) with details such as the date, time, country, number of records, and the deviation from average. Furthermore, a visual summary of the pattern's total records is shown in a horizontal bullet chart. For example, selecting four different countries results in very different patterns in the heatmap, as shown in Figure 3. We provide a video showcasing all possible interactions in Supplemental Materials.

All interactions with the dashboard require one click or less, meaning a user can hover over any element for an updated view of the patterns in BubbleNet. This hover over affects all other views, and BubbleNet also provides a pop-up of this selection as in Figure 2a). By clicking on any element, that selection becomes locked in place and updates the selection window in Figure 2e). Otherwise when a user hovers off an item, then its previous selection is reset.

By default, the initial pattern is the most recent day and the entire world. As such, the bullet charts in Figure 2c) look like regular bar charts until a country is selected to show this country as a subset of the world's pattern. Through feedback from users, it was found that reducing clicks for selection was desired in a dashboard setting and enabled fast comparison of two selections, by selecting one element and hovering on and off another element. Keyboard interactions were also added to more easily navigate selections through time and to reset back to the default pattern.

One can compare the interaction of each view with our tasks, back in Section 3. For location, temporal, and attribute views, all of the elements were interactive, e.g. hovering or clicking on a country, day, hour, or type of attribute. These selections supported pivoting data to identify and compare patterns. The records table view supports identification and comparison of patterns but not pivoting since analysts often use their own tools for this purpose.

5.6. Implementation

The BubbleNet dashboard presented in Figure 2 was created using D3.js for all visualization components. Each interaction filters a different portion of the same dataset loaded in the web browser. These datasets are prepared via a set of back-end Python scripts which aggregate network security datasets into summaries by day, broken apart by location and by hour with statistics pre-computed on the data. Lastly, these daily summaries are combined in Python to produce JSON files for the web dashboard, so real-time data is possible but currently requires a refresh of the page.

The visualizations shown in this paper, in the video, and included in the usability study all showcase real data from a large organizational network, capturing a summary of a month's worth of data or about a million records. In particular, the dataset shown is from an intrusion detection system, which automatically flags important network records as alerts for network analysts. These alerts can be generated by pre-defined rules, which is most often the case, or by more sophisticated machine learning techniques. The BubbleNet dashboard is designed in such a way to support visualization of any dataset which can be broken into network records and geolocated, so it works best when analyzing traffic over the internet. When it comes to scalability, the dashboard maintains interactivity with millions of records due to aggregation done on the back-end.

6. Evaluation and Deployment

Evaluation is undoubtedly an important aspect to designing tools for users, both for cyber security [SYC*14] but more broadly as well. First, we discuss the evaluation methodology of a usability study. This study is a combination of formative and summative evaluation since key issues were prioritized on a high-fidelity prototype but user needs were also uncovered. The results of this study highlight the usability of BubbleNet, and the BubbleNet dashboard in Figure 2 was thus deployed in a research environment. However, this study also highlighted missing elements of utility from the BubbleNet dashboard, so a final design iteration was required to address these elements and deploy the tool in an operational environment.

6.1. Evaluation Methodology

To improve upon the second prototype from Figure 1b), a usability study was performed with network analysts and managers from both research and operational organizations using real-time, real-world data from an organizational network. The intent of this study was to improve the design and see if the prototype met the needs of both analysts and managers. Nine cyber security professionals participated in the study: five analysts, four managers. Each participant took part in a one-hour long think-aloud session, conducted by one moderator with an observer taking notes, both of whom are co-authors on this paper. Each session contained a scripted walkthrough of the prototype, several prescribed tasks to complete, open-ended questions about how users would use the prototype, and distribution of a system usability scale [Bro96].

To analyze data from the think-aloud session, the notes taken by the observer were analyzed with a qualitative coding methodology [SC90]. This coding was conducted by the primary author, through an open tagging of two users' comments and consolidating tags to all other user comments. Furthermore, the system usability scale is a standardized survey technique [Bro96] used to evaluate the prototype's usability, and other researchers have utilized such a survey [LN13]. This usability survey has been used to evaluate the usability of systems for 30 years with its set of 10 standardized statements rated on a Likert scale, and it works well with a small group of users [Sau11]. By combining this survey with a qualitative coding methodology, we sought to increase the analytical rigor of evaluating our prototype to determine if it was ready to be deployed to users.

6.2. Evaluation Results

After coding each of the participants' comments, the following categories of tags were formed: desired task, that task's intended target in the dashboard, and its outcome. Example tasks include to present, filter, or identify with any of the views presented in Section 5, and example outcomes include successes, struggles, and failures along with other tags such as feature suggestions. These tags provided a unique view on the qualitative data, and a list of features were prioritized and implemented for BubbleNet in Figure 2. These features that were added include: details-on-demand records view, better selection feedback, new map color encoding, and keyboard interactions. This analysis process gathered the key successes of the BubbleNet dashboard:

- Temporal pattern detection was simple and easy using the heatmap: *"I keep getting drawn to the heatmap and these darker areas, because they certainly stand out"* (A4) & *"[heatmap] helps find those temporal patterns"* (A1)
- Users expressed that the dashboard's utility was for discovering patterns and trends in the data: *"the majority of what we are looking for is patterns and this just makes patterns which is faster"* (A4)
- One-click-or-less interactions worked very well: *"it's very responsive and dynamic; the fact that it changes as I narrow [in] is the best"* (M2)
- Most interactions occurred with the bullet charts and heatmap: *"I could write a splunk query to do this, but this is easier"* (A5)
- No expressed dislike for animation in the map view: *"best part is the instant visual gratification"* (A4)

Furthermore, this analysis derived a set of design considerations for future cyber security dashboards, presented in Section 7. With the first few participants of the study, a common usability issue was discovered since the bullet charts had two different bars to click on. Along with visual bugs, these issues were fixed right away to focus feedback on less obvious issues. Quantitatively, this can cause issues, but, since the changes were motivated by and reduced user frustrations, we hypothesize that the quantitative results from the usability survey would have only improved if we had re-run the study with these fixed usability issues.

The prototype gave users novel insights on their data. For example, one participant found a pattern in a particular country and told us that they *"never would have got[ten] there by looking at the alerts in text format"* (A1). This same analyst told us that they could imagine this dashboard being used with other kinds of datasets as well: *"pretty much everything: flow data, [firewall logs], [proxy logs], anything"* (A1). This statement helped confirm that the abstraction was at just the right level since the dashboard could adapt to so many cyber security datasets.

We present the quantitative results of the usability survey in Figure 4. The system usability scale provides a standard set of questions where an average system would receive a score of 68 out of 100 [Sau11], and the usability of our prototype was found to be above average: 74.7. We provide the data and results from the usability survey in Supplemental Materials. Each individual question can be broken into a set of characteristics [Sau11], and by doing so we found that the BubbleNet dashboard scored high on learnability and ease of use. By analyzing the results of analysts versus managers, we found no significant differences. However, network managers rated BubbleNet as less complex, less cumbersome, and easier to learn. We did have one outlier (A8), who was two standard deviations lower than the average, which lowered the final score due to the relatively small sample size. We hypothesize that this user simply rates things more strictly since this user still achieved tasks successfully and had similar concerns as other analysts.

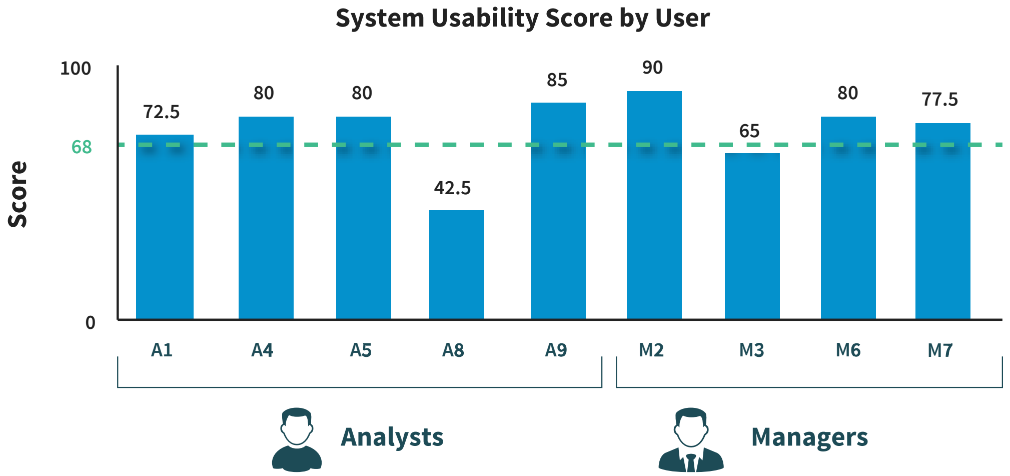


Figure 4: *Final results of a system usability survey of nine different users, both network analysts and managers. The average score of the dashboard is 75, above the average usability score of 68 [Sau11].*

6.3. Deployment

After the usability study, further development led to the final BubbleNet dashboard. Then, BubbleNet was deployed to users with real-time data in a research network operations center. However, BubbleNet was developed and deployed with only a single data source and a short time range, so it was arguable how useful its design could be for other users. This is coupled with the fact that the usability survey scored lower on a question that arguably could be interpreted with respect to its utility: *"I think that I would like to use this system frequently."*

To gauge its operational utility, the BubbleNet dashboard was further demonstrated with multiple relevant datasets to different analysts at three cyber operations centers. Analysts and managers provided qualitative feedback via comments, both during the demonstration as a group, and in private conversations afterwards. These demonstrations, feedback, and design iterations took place in the fourth design channel of Figure 1. In summary, this feedback highlighted the simplicity of the flat map, conjunction of small multiples with interaction, and a critical area for improvement with respect to scaling to multiple data sources.

This feedback from operational analysts led to the final design iteration and deployed operational tool. To incorporate multiple data sources, significant trade-offs existed between displaying all data and the tight integration required for linked small multiples as presented in BubbleNet. As such, this final tool utilizes the assembly-canvas metaphor [Ogr09], similar to Tableau's dashboards where a custom visualization dashboard is built on the fly. The flat map serves as the background for any geospatial data. There is a leftmost palette which lists the available data sources. When selecting data sources that are not geospatial, a floating visualization palette is placed on the screen for the user to select a different visualization for the data. These palettes support customization of numerous visualizations: e.g. treemap, node-link diagrams, sunburst charts, and timelines, and this customized dashboard can be saved and shared.

After implementing this final tool, end users have expressed an interest in adopting for daily use. Next steps for the project include a formal, summative end user evaluation. While developing this final design, we identified several design considerations for future development, such as establishing consistent visual encodings across varied datasets and connecting these visualizations through interaction. While out of scope for this project, these considerations remain important for continuing operational deployment.

7. Reflections

We uncovered a set of implications for dashboard visualization of cyber security data which others can use. First, analysts want details of the data whereas managers sought the broader impact of an incident on the larger network. Secondly, there are many different ways to aggregate and provide details of the underlying data, so it is imperative to use and adapt multiple cyber visualizations to different needs over time. Third, it was discovered that a map for cyber data is not completely useless. Users are able to situate themselves and pivot data to find novel insights, and a map is one way to scaffold a visualization into other kinds [Mar15]. Fourth, fast hover-over interactions are very appropriate to reduce the number of required clicks to pivot visualizations using animation and provide quick details-on-demand.

Upon reflecting on this design study, we realized that winnowing and casting of user roles [SMM12] occurred later in the user-centered design process highlighted in Figure 1. Unlike a typical design study, there was very limited time from a single set of domain experts. By reviewing previous detailed cognitive studies of users and through interviews, personas were crafted to identify different potential users [MSM15]. As a result, users were winnowed into two types, analysts and managers. This approach was motivated by domain constraints: limited access to users and data.

Another reason behind this unique design process is due to the task of presentation. Presentation inherently involves two or more parties, so it could involve users beyond a data analyst. In a design study methodology, Sedlmair et al. describe several different kind of collaborator roles, such as front-line analysts and gatekeepers [SMM12]. Alternative collaborator roles have been identified, such as liaisons [SMKS15] which bridge visualization research to complex domains. While we worked with several liaisons, the user personas identified four kinds of users where only one, the network analyst, is a domain expert in cyber security. Other users, such as network managers, have some domain knowledge, but there was clearly another domain at work here: an organizational domain. Large organizations need to disseminate information up a chain of command in order for decisions to be made and passed down [MSM15]. With multiple domains and types of users, this work challenges the role of a single domain expert as the optimum collaborator. It is important to identify these different user roles and design tools which adapt to their needs.

Lastly, working in the cyber security domain has benefited from the multiple discourse channel approach [WBD14] as highlighted in Figure 1. By reflecting on our design process, this multiple channel approach is particularly beneficial with the unique design constraints we faced: limited access to users and data, multiple types of users, and balancing trade-offs to deploy tools. The design of BubbleNet occurred within the second channel at a research organization, but this design would not have been as successful without the design methods and knowledge gained from the other channels. For example, the third channel represents a collaboration with a university network analyst which enabled us to validate abstractions of network security data and critically changed BubbleNet's location view. By working at an operational organization in the fourth channel, BubbleNet's design influenced and inspired new encodings to be implemented by a team of developers, leading to operational tool deployments. As discussed in Section 6, deploying a tool is a complex process which involves further design trade-offs, but it is important to discuss these aspects to help further the field of visualization and get tools in the hands of users.

8. Conclusion

In this work, we have presented the first complete visualization design study for cyber security, resulting in a novel, interactive real-time dashboard which was deployed in both research and operational environments. This design study involved multiple projects and various user-centered design methods to achieve these goals. The work presented here is part of an ongoing investigation in order to overcome general challenges such as limited access to users and data.

However, the BubbleNet dashboard is not the end of research or development into cyber security dashboards. The use of a map does not work for all data, and there is more work needed to find more effective encodings such as broader impact of cyber security incidents. Nevertheless, the design process of BubbleNet shows how other design studies can work with collaborators and users beyond just data analysts. When working with these other types of users, it becomes more important to balance and prioritize appropriate sets of user needs to design, develop, and deploy effective, domain-specific visualization tools.

9. Acknowledgments

The authors wish to thank Jonzy, Dan Bowden, and Tamara Denning for the data sketches method, staff members at MIT Lincoln Laboratory for the personas method, and the Visualization Design Lab at the University of Utah for their feedback on this work. This work is sponsored in part by the Air Force Research Laboratory and the DARPA XDATA program. The Lincoln Laboratory portion of this work was sponsored by the Assistant Secretary of Defense for Research & Engineering under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

References

[ALK*11] AMERSHI S., LEE B., KAPOOR A., MAHAJAN R., CHRISTIAN B.: CueT: Human-guided fast and accurate network alarm triage. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11* (New York, New York, USA, may 2011), ACM Press, p. 157. 2

[BDF*08] BLUE R., DUNNE C., FUCHS A., KING K., SCHULMAN A.: Visualizing real-time network resource usage. In *VizSec* (Berlin, Heidelberg, Sept. 2008), Goodall J. R., Conti G., Ma K.-L., (Eds.), vol. 5210 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 119–135. 2

[BDJ14] BARGO R., DEARDEN J., JONES M. W.: Order of magnitude markers: An empirical study on large magnitude number detection. *IEEE Trans. on Visualiz. and Comp. Graphics 20*, 12 (Dec. 2014), 2261–2270. 5

[BEK14] BEST D. M., ENDERT A., KIDWELL D.: 7 key challenges for visualization in cyber network defense. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (New York, NY, Nov. 2014), ACM Press, pp. 33–40. 1

[Bos15] BOSTOCK M.: Pseudo-Dorling Cartogram. http://bl.ocks.org/mbostock/4055892, 2015. 4

[Bro96] BROOKE J.: Sus-a quick and dirty usability scale. *Usability evaluation in industry 189*, 194 (1996), 4–7. 2, 7

[CLK09] CHOI H., LEE H., KIM H.: Fast detection and visualization of network attacks on parallel coordinates. *Computers & Security 28*, 5 (July 2009), 276–288. 2

[CSL*16] CAO N., SHI C., LIN S., LU J., LIN Y.-R., LIN C.-Y.: TargetVue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (jan 2016), 280–9. 2

[Dor11] DORLING D.: *Area Cartograms: Their Use and Creation*. John Wiley & Sons, Ltd, Chichester, UK, Apr. 2011. 4

[DW08] D'AMICO A., WHITLEY K.: The real work of computer network defense analysts. *Proceedings of the Workshop on Visualiz. for Cyber Sec.* (2008), 19–37. 1, 2, 3

[ECS05] ERBACHER R., CHRISTENSEN K., SUNDBERG A.: Designing visualization capabilities for IDS challenges. In *Workshop on Visualiz. for Comp. Sec.* (2005), pp. 121–127. 2, 3

[EFC*10] ERBACHER R. F., FRINCKE D. A., CHUNG WONG P., MOODY S., FINK G.: A multi-phase network situational awareness cognitive task analysis. *Information Visualization 9*, 3 (Jan. 2010), 204–219. 1

[EFW*10] ERBACHER R. F., FRINCKE D. A., WONG P. C., MOODY S., FINK G.: Cognitive task analysis of network analysts and managers for network situational awareness. In *IS&T/SPIE Electronic Imaging* (2010), pp. 75300H–75300H. 2

[FAL*06] FORESTI S., AGUTTER J., LIVNAT Y., MOON S., ERBACHER R.: Visual correlation of network alerts. *IEEE Computer Graphics and Applications 26*, 2 (Mar. 2006), 48–59. 2

[Few10] FEW S.: Bullet graph design specification. *Perceptual Edge-White Paper* (2010). 5

[FJS*14] FOWLER J. J., JOHNSON T., SIMONETTO P., SCHNEIDER M., ACEDO C., KOBOUROV S., LAZOS L.: IMap: Visualizing network activity over internet maps. *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (2014). 2, 3

[FK14] FISCHER F., KEIM D.: NStreamAware: Real-time visual analytics for data streams to enhance situational awareness. *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (2014). 2, 3

[FMN05] FINK G., MUESSIG P., NORTH C.: Visual correlation of host processes and network traffic. In *IEEE Workshop on Visualiz. for Comp. Sec.* (2005), IEEE, pp. 11–19. 2

[FNER09] FINK G. A., NORTH C. L., ENDERT A., ROSE S. J.: Visualizing cyber security: Usable workspaces. In *Proceedings of the Workshop on Visualiz. for Cyber Sec.* (2009), pp. 45–56. 1

[GS09] GOODALL J. R., SOWUL M.: VIAssist: Visual analytics for cyber defense. In *2009 IEEE Conference on Technologies for Homeland Security* (May 2009), IEEE, pp. 143–150. 2, 3

[HHH13] HAO L., HEALEY C. G., HUTCHINSON S. E.: Flexible web visualization for alert-based network security analytics. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (New York, NY, Oct. 2013), ACM Press, pp. 33–40. 2

[HL98] HOWARD J. D., LONGSTAFF T. A.: A common language for computer security incidents. *Sandia National Lab.* (1998). 2

[Kas15] KASHAN O.: information is beautiful: World's Biggest Data Breaches. http://informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/, 2015. 1

[KOK05] KOIKE H., OHNO K., KOIZUMI K.: Visualizing cyber attacks using IP matrix. In *Proceedings of the Workshop on Visualiz. for Cyber Sec.* (2005), IEEE, pp. 91–98. 2, 3

[KRA*05] KOMLODI A., RHEINGANS P., AYACHIT U., GOODALL J., JOSHI A.: A user-centered look at glyph-based security visualization. In *Proceedings of the Workshop on Visualiz. for Cyber Sec.* (2005), IEEE, pp. 21–28. 2, 3

[LD11] LLOYD D., DYKES J.: Human-centered approaches in geovisualization design: Investigating multiple methods through long-term case study. *IEEE Trans. on Visualiz. and Comp. Graphics 17*, 12 (2011), 2498–2507. 3

[LHS14] LANDSTORFER J., HERRMANN I., STANGE J.: Weaving a carpet from log entries: A network security visualization built with co-creation. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology* (2014). 2, 3

[LN13] LESCHKE T. R., NICHOLAS C.: Change-link 2.0: A digital forensic tool for visualizing changes to shadow volume data. In *Proceedings of the Workshop on Visualiz. for Cyber Sec.* (New York, NY, Oct. 2013), ACM Press, pp. 17–24. 2, 7

[LYL04] LAKKARAJU K., YURCIK W., LEE A. J.: NVisionIP: netflow visualizations of system state for security situational awareness. In *Proceedings of the ACM Workshop on Visualization and Data Mining for Computer Security* (New York, NY, Oct. 2004), ACM Press, p. 65. 2, 3

[Mar15] MARAI G. E.: Visual Scaffolding in Integrated Spatial and Non-spatial Analysis. The Eurographics Association. 8

[Max15] MAXMIND: GeoLite2 Free Databases. http://dev.maxmind.com/geoip/geoip2/geolite2/, 2015. 3, 4

[MMAM14] MCKENNA S., MAZUR D., AGUTTER J., MEYER M.: Design activity framework for visualization design. *IEEE Trans. on Visualiz. and Comp. Graphics 20*, 12 (2014), 2191–2200. 3

[MP08] MUSA S., PARISH D. J.: Using time series 3D alertgraph and false alert classification to analyse snort alerts. In *Visualiz. for Comp. Sec.* (2008), vol. 5210, Springer, pp. 169–180. 2

[MSM15] MCKENNA S., STAHELI D., MEYER M.: Unlocking user-centered design methods for building cyber security visualizations. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (2015), IEEE. 1, 2, 3, 8, 9

[Ogr09] OGRINZ M.: *Mashup patterns: Designs and examples for the modern enterprise*. Pearson Education, 2009. 8

[PRS*13] PAUL C., ROHRER R., SPONAUGLE P., HUSTON J., NEBESH B.: CyberSAVI: A cyber situation awareness visual interface for mission-level network situation awareness. *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (2013). 2

[RER*10] RASMUSSEN J., EHRLICH K., ROSS S., KIRK S., GRUEN D., PATTERSON J.: Nimble cybersecurity incident management through visualization and defensible recommendations. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (New York, NY, Sept. 2010), ACM Press, pp. 102–113. 2

[Sau11] SAURO J.: Measuring Usability with the System Usability Scale. http://measuringu.com/sus, 2011. 7, 8

[SC90] STRAUSS A., CORBIN J.: *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. 1990. 7

[SMKS15] SIMON S., MITTELSTÄDT S., KEIM D. A., SEDLMAIR M.: Bridging the gap of domain and visualization experts with a Liaison. *Eurographics Conference on Visualiz.* (2015). 9

[SMM12] SEDLMAIR M., MEYER M., MUNZNER T.: Design study methodology: Reflections from the trenches and the stacks. *IEEE Trans. on Visualiz. and Comp. Graphics 18*, 12 (2012), 2431–2440. 1, 8, 9

[SYC*14] STAHELI D., YU T., CROUSER R. J., GWYNN D. O., MCKENNA S., HARRISON L.: Visualization evaluation for cyber security : Trends and future directions. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (2014), pp. 49–56. 2, 7

[TPG*09] TAYLOR T., PATERSON D., GLANFIELD J., GATES C., BROOKS S., MCHUGH J.: FloVis: Flow visualization system. In *Cybersecurity Applications & Technology Conference for Homeland Security* (Mar. 2009), IEEE, pp. 186–198. 2

[Wat05] WATSON B.: IDGraphs: Intrusion detection and analysis using histographs. In *Proceedings of the Workshop on Visualiz. for Comp. Sec.* (2005), IEEE, pp. 39–46. 2, 3

[WBD14] WOOD J., BEECHAM R., DYKES J.: Moving beyond sequential design: Reflections on a rich multi-channel approach to data visualization. *IEEE Trans. on Visualiz. and Comp. Graphics PP*, 99 (2014), 1–1. 1, 3, 9

[WD08] WOOD J., DYKES J.: Spatially ordered treemaps. *IEEE Trans. on Visualiz. and Comp. Graphics* (2008). 3

[YLRB10] YU T., LIPPMANN R., RIORDAN J., BOYER S.: EMBER: A global perspective on extreme malicious behavior. In *Proceedings of the Symp. on Visualiz. for Cyber Sec.* (New York, NY, Sept. 2010), ACM Press, pp. 1–12. 3

[YYT*04] YIN X., YURCIK W., TREASTER M., LI Y., LAKKARAJU K.: VisFlowConnect: Netflow visualizations of link relationships for security situational awareness. In *Proceedings of the Workshop on Visualiz. and Data Mining for Comp. Sec.* (New York, NY, Oct. 2004), ACM Press, p. 26. 2

[ZCW*14] ZHAO J., CAO N., WEN Z., SONG Y., LIN Y.-R., COLLINS C.: #FluxFlow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (dec 2014), 1773–82. 2