EEVi –Framework and Guidelines to Evaluate the Effectiveness of Cyber-Security Visualization

Aneesha Sethi, Federica Paci, Gary Wills
Electronics and Computer Science, University of Southampton, UK

Abstract

Cyber-security visualization aims to reduce security analysts' workload by presenting information as visual analytics instead of a string of text and characters. However, the adoption of the resultant visualizations by security analysts, is not widespread. The literature indicates a lack of guidelines and standardized evaluation techniques for effective visualization in cyber-security, as a reason for the low adoption rate. Consequently, this article addresses the research gap by introducing a framework called EEVi for effective cyber-security visualizations for the performed task. The term 'effective visualization' is defined as the features of visualization that are critical for an analyst to competently perform a certain task. EEVi has been developed by analyzing qualitative data which led to the formation of cognitive relationships (called links) between data. These relationships acted as guidelines for effective cyber-security visualization to perform tasks. The methodology to develop this framework can be applied to other fields to understand cognitive relationships between data. Additionally, the analysis of the framework presented, demonstrates how EEVi can be put into practice using the guidelines for effective cyber-security visualization. The guidelines can be used to guide visualization developers to create effective visualizations for security analysts based on their requirements.

1. Introduction

In the field of cyber-security, public and private sectors rely on the expertise and capabilities of security analysts to protect assets and resources connected via computer networks. An area under the umbrella of cyber-security is cyber-security visualization, which provides these analysts with visual data rather than textual data for analysis. Cyber-security visualization aims to provide effective tools [1] that help detect, monitor and mitigate sophisticated technical and social attacks in a timely manner. Thus, it focuses on providing security analysts with a competent weapon to prevent and defend against cyber-attacks.

There is an outburst of cyber-security visualization tools, but the visualizations presented by these tools are rarely evaluated for effectiveness in terms of the task they aid in performing [2]. Moreover, most of the visualizations are developed and sometimes evaluated without any involvement of users [2]. The lack of user-involvement could be a result of limited access to security analysts; or due to the nature of security analysts' jobs they cannot make significant time-commitments [3]. This leads to low adoption rate of such tools [4].

This article introduces EEVi, a framework for evaluating the effectiveness of visualization in cyber-security. EEVi was developed by qualitatively analyzing requirements of security analysts based on the task they perform, and using the resultant data to form cognitive relationships that form fundamental guidelines. These guidelines are used to guide developers towards creating cyber-security visualizations that are effective for a security analyst performing a specific task.

In this article, Section II introduces the background literature that led to the identification of the research gap. Section III describes the methodology of qualitative coding that was followed to develop the framework and Section IV introduces the structure of the framework developed as a result. Section V presents the analysis conducted from the framework that can be used as fundamental guidelines for cyber-security visualization. Finally, Section VI concludes this article with an overview behind the rationale of this research.

2. Background Literature

The authors discovered a research gap while reviewing the literature; most cyber-security visualization tools introduced had minimal or no evaluation of the visualizations that were presented. Consequently, a user could not judge the effectiveness of these visualizations for the tasks performed by security analysts. The following section explains the background literature that led to and aided in the development of the framework.

In the field of cyber-security there is an explosion of tools that focus on different aspects of cyber-security visualization ranging from a high-level view of the system to a technical low level view. Most of these tools can be broadly classified into three categories: network analysis, malware analysis and insider threat analysis.

Network Analysis tools focus on mapping the physical network of the system to detect possibilities of attack. It includes tools which visually monitor network traffic using intrusion detection techniques like in *CyberVis* [5] or proactive tools that display graphs to highlight potential attack vectors based on the state of the network like *PERCIVAL* [6]. All of these tools use different kinds of visualizations ranging from the attack-graphs [6] to complicated customized visualizations.

Malware Analysis tools focus on identifying, detecting and eliminating malware. It includes tools that focus on visually detecting rogue autonomous systems indicating possible malware like *BURN* [7] or tools that detect malware attacks and determine its effects like *DAVAST* [8]. These tools mainly use different kinds of graphs and charts to present the analysis.

Insider Threat Analysis focuses on analyzing attacks by malicious insiders, people who intentionally try to misuse the legitimate information they have access to. It includes tools that visually detect anomalies and possible attacks through pattern matching [9] or by using machine learning to check for anomalous behavior [10]. These tools use visualizations like color maps [9] or different types of graphs like attack-pattern trees [10]

Most of the aforementioned tools provide situational awareness. Situational awareness is a high-level abstract view [11] of the system which presents an overview and is beneficial to both technical and non-technical people as it aims to bridge the knowledge gap between the two. However, these tools have not been evaluated to determine their effectiveness in terms of the task they support in performing. Staheli et al. [2] presented a survey, in 2014, of 130 VizSec (IEEE Symposium on Visualization for Cyber Security) papers which showed that little research had been conducted in determining the effectiveness of cyber-security visualization. It also showed that 46% of these papers did not have any user-involvement in the evaluation phase. To reinforce the results of the survey by Staheli et al. [2], the authors conducted a survey of nine papers. It was observed that two of these had no form of evaluation, three did not have any user-involvement and only one allowed complete and unguided user-interaction with the visualization. Additionally, there was a lack of standardization among the evaluation techniques used to evaluate these nine tools.

Therefore, the resulting visualizations presented by cyber-security visualization tools were not effective and usually did not consider the needs of the user or involve them in the evaluation or formative processes. This resulted in a low adoption rate of these tools [4]. Additionally, the evaluation techniques used to evaluate most tools were not effective in evaluating the visualizations that were produced based on the performed task nor were the evaluation techniques standardized. The lack of a common framework for standardized evaluation methods [2] has been highlighted repeatedly within the literature. However, there is no research supporting the development of such a framework to evaluate the effectiveness of cyber-security visualization tools based on user requirements. An assessment of user requirements must be included during the early design phases and later evaluation phases. Thus, creating a need for guidelines to standardize evaluation techniques and utilize them to evaluate for effectiveness of the performed task.

The main challenge faced in conducting research to develop such a framework in this area, is the lack of access to experts. This was overcome by using cognitive task analysis (CTA) papers. The idea of using CTA papers was initially introduced by Mckenna et al. [3], who used qualitative coding of CTA papers to form requirements for the cyber-security visualization tool they were developing. One of the goals of CTA analysis conducted by D'Amico et al. [12] was having the resultant analysis used as foundation material for studies with lack of access to security experts or analysts. Additionally, qualitative coding was used by Lam et al. [13] to describe different evaluation techniques currently used to evaluate visualizations. This led to a need for cognitive task analysis (CTA) papers for cyber-security visualization to develop the framework.

Vessey's theory of cognitive fit has a classification of spatial tasks, which require problems to be looked at as a whole and require "...making associations or perceiving relationships in the data" [14] to find effective solutions for the problems. Thus, EEVi is developed on analysis of qualitative papers to form cognitive relationships for effective guidelines for cyber-security visualization. More details about the methodology and process of qualitative coding is explained in the next section.

3. Methodology

EEVi was developed using Thematic Analysis which is a qualitative bottom-up approach. A bottom-up approach means going through the data,

without any pre-conceived notions, to completely develop notions or in this case, themes and codes. Thematic Analysis is one such qualitative analytic method used to identify, examine and report patterns (or themes) within data [15]. The four major steps of Thematic Analysis (see Figure 1) used to develop the framework EEVi [15] are explained in the following sub-sections.

3.1 Familiarizing with data

The data that formed the basis of EEVi was mainly derived from papers that presented results of Cognitive Task Analysis (CTA) of security analysts. CTA attempts to follow an inductive approach rather than trying to identify predefined data [16]. It has been used in many studies to describe the cognition (or the way the mind works) necessary for task performance and to extract mental models or in this case, how analysts achieve situational awareness for cyber-security [17]. Most studies generally include interviews, observations and hypothetical scenario creation [12].

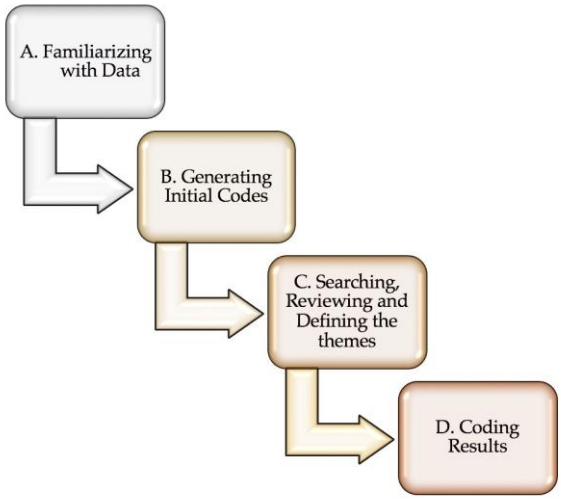

Figure 1. Overview of the methodology followed for Thematic Analysis.

The papers that were used for the purpose of this article were selected because of the data they represented. These papers gave precise details about analyst roles, the type of data they used, how the analyses were conducted, what the analysts thought about visualization approaches and their experiences, if any, with visualizations. D'Amico et al. [12] and D'Amico et al. [18] gave insight into roles of analysts and the tasks they perform in organizations. Erbacher et al. [11] presents interviews with analysts for the specific purpose of cyber-security visualization. Fink et al. [1] presents a variety of information about how to make visualizations effective for security analysts (who were the end users) and Mckenna et al. [3] formed the basis of this study and helped understand how to research these papers and take out the relevant elements from them. These papers formed the background for EEVi.

3.2 Generating initial codes

The next step began after the papers were studied for an overall understanding of the data [19] and an initial list of ideas represented in the papers had been formulated. This list was further analysed, at this stage, and was used to produce the initial set of codes. The process of coding reduces the amount of raw data by breaking down data to manageable high-level abstractions called codes [19]. These codes represent an excerpt of data and are used to intuitively identify the aspects of the data it represents. The codes represent the qualitative aspects of the framework which is represented by a codebook which includes a collated list of all the initially generated codes. At this stage an idea of the themes start to form but are not yet defined.

3.3 Searching, reviewing and defining the themes

The next step is to define the themes based on the codebook that has been generated in the previous step. A theme captures the significance of the data and represents a patterned response [6] which is reflected by the group of codes it defines.

The codebook consists of a list of different codes that were identified across the dataset. The next step was to organise the codes and compare them to find the similarities and differences, a potential theme was attached to each cluster of similar codes [19]. At this stage the relationships between the potential themes and codes also starts to form.

The potential themes are then reviewed against the codes they represent and further refined. Then the potential themes are reviewed against the data, literature and research questions to validate the representation of the theme by reviewing the relationships the themes form against the data.

The themes are therefore defined according to the data they represent and how they fit in relation to the data set and the research questions. The identified themes were:

1. Analysis of Data – Task performed by security analysts;
2. Data – Type of data used to perform tasks;
3. Feature of Visualization – Features of visualization required to perform the tasks;

4. Role of Analyst – The security analyst that perform the tasks.

3.4 Coding results

The results identify the themes, codes and relationships or links identified as a result of thematic analysis of the dataset. The cognitive relationships, defined as links, influenced the development of EEVi by linking the themes into the model of the framework. The cognitive relationships formed between different codes led to a similar generic storyline of themes. A storyline presents the narrative of a coherent story through which themes can be described and cognitively linked [19]. This storyline was defined and formed the structure of EEVi.

4. Structure of EEVi

The results identified in the previous section were cognitively linked together and led to EEVi to determine the effectiveness of visualization depending on the performed task. The structure of the framework can be seen in Figure 2.

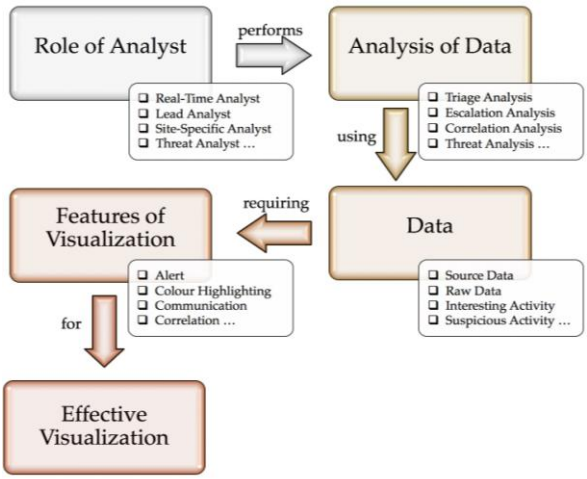

Figure 2. Structure of EEVi for deriving effective visualization

The framework forms links (or cognitive relationships) between the four themes to determine the effectiveness of a tool. These themes are defined by a set of codes. When these codes are cognitively linked together with the data from the qualitative analysis, it forms a cognitive relationship, which is represented by storylines. These storylines identify relevant codes from themes that would be imperative for an effective visualization for the performed task when linked together cognitively. Thus, the generic structure of the framework can be defined as Role of Analyst *performs* Analysis of Data *using* Data and *requiring* Features of Visualization *to create* effective cyber-security visualization. The resultant visualizations would be effective as they would include the features that are required by security analysts to perform a task competently, as demonstrated in the following section.

5. Analysis of EEVi – Guidelines for effective cyber-security visualization

The structure of EEVi was developed using a storyline which led to the generic framework. The guidelines of effective visualization for cyber-security based on this framework can be defined by tasks performed by security analysts.

The framework represents cognitive relationships for each task to determine the critical features of visualization that are required to allow the security analyst to competently perform the task. The cognitive relationships represent guidelines for effective visualization of each task identified. Therefore, the term *effective visualization* is defined as features of visualization that are critical for a security analyst to competently perform a specific task. These features of visualization represent the resources required by a security analyst to perform a task effectively and not the esthetics (like the type or color of graphs) that would be required by a security analyst. The features can be identified in the guidelines represented in the following sub-sections. The purpose of these guidelines is to guide visualization developers towards creating effective visualizations for cyber-security. Thus, reducing the knowledge gap between security analysts (end-users) and visualization developers.

The results of the qualitative coding led to the identification of eight such tasks. These eight tasks were identified from the data gathered from CTA papers as they were the most common tasks conducted by security analysts. These tasks are identified by different names in different organizations [12] but, they are performed in every organization. The purpose of each task is clear from the definitions.

The eight tasks that were identified from the data analyzed during the qualitative coding process are discussed in the following sub-sections. These are represented along with the definition of the task, the identification of the analyst who performs the task, the data used to perform the task and the features of visualization that would lead to an effective visualization. These are explained with the codes and excerpts of data represented by the codes, to demonstrate the logic that led to the development of the cognitive relationships, and hence the guidelines.

5.1 Guidelines for effective visualization of triage analysis

Triage Analysis is the first look at data [12]. At this stage the analyst weeds out false positives for further analysis [18]. It is performed within an order of a few minutes [11].

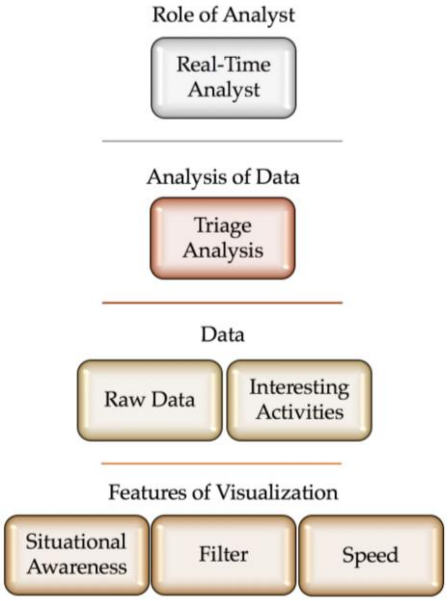

Figure 3. Visual representation of cognitive relationship for Triage Analysis, as derived from the framework.

A visual representation of this relationship is displayed in Figure 3. The excerpts and codes that led to this relationship are explained as follows: Triage Analysis is usually performed by a Real-Time Analyst [12]. It is the "...first look at the raw data and interesting activity" [12] and hence uses Raw Data and Interesting Activities as types of Data. Raw Data is the most elemental data, usually in very large quantity and is passed through automated process to filter and Interesting Activities is data that has been flagged by automated processes on raw data and is inspected by an analyst, usually contains a large number of false positives [12]. Visualization for Triage Analysis requires abilities to Filter for "...initial filtering" [12] and for "...weeding out false positives..." [18]. Filter allows the ability to easily filter, join or transform data without changing the original [1]; and also, allows ability to filter noise to allow analyst to see trends [11]. It also requires Speed of data access as "...triage period should be on the order of minutes" [11] and a "...relatively fast decision..." [12] needs to be made. Another important feature for Triage Analysis is having Situational Awareness as triage is performed at "...a highly abstract, situational-awareness level" [11]. Situational Awareness gives an accurate picture of external and internal information in an overview to allow for rapid decision making and to allow for analysts to understand the state of all resources [11]. This would lead to an effective visualization for a Real-Time Analyst performing Triage Analysis.

5.2 Guidelines for effective visualization of escalation analysis

Escalation Analysis is investigation of suspicious activities from the previous stage and production of reports [12]. It may take from hours to a few weeks to complete [18].

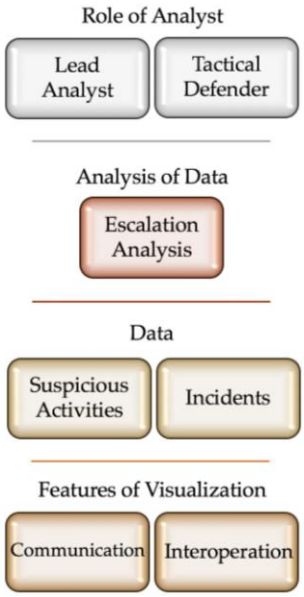

Figure 4. Visual representation of cognitive relationship for Escalation Analysis, as derived from the framework

A visual representation of this relationship is displayed in Figure 4. The excerpts and codes that led to this relationship are explained as follows: Escalation Analysis is usually performed by Lead Analyst [12] along with a Tactical Defender [1]. A Tactical Defender defends against current and immediate attacks [12] by maintaining situational awareness of the system and rapid rectification of problems [1]. They "...investigate suspicious activity[ies]" [12] and hence uses Suspicious Activity as a type of Data. Suspicious Activities is data that is anomalous after the initial triage analysis and needs to be monitored [12]. It also uses Incidents as a "...goal of escalation analysis is to produce incident reports" [12] as the type of Data. Incidents are defined at the point when the occurrence and seriousness of an event is confirmed and formally reported [12]. Visualization for Escalation Analysis requires Communication as it is based on "...tip-offs

from colleagues and cooperating organizations" [18]. Communication enables users to communicate and collaborate with other analysts [11] by sharing findings [1, 3] and providing support for report building [12]. It also requires Interoperation of data as "...the analyst marshals more data, usually from multiple data sources..." [18], interoperations are the ability of a tool to work efficiently with other tools, applications, utilities or databases [1]. This would lead to an effective visualization for a Lead Analyst or Tactical Defender performing Escalation Analysis.

5.3 Guidelines for effective visualization of correlation analysis

Correlation Analysis is the search for patterns and trends in data, which may be previously unrecognized [12, 18].

A visual representation of this relationship is displayed in Figure 5. The excerpts and codes that led to this relationship are explained as follows: Correlation Analysis is performed by Site-Specific Analyst [12] along with a Tactical Defender [1]. It "...includes grouping data into intrusion sets" [12] and hence uses Intrusion Sets as type of Data. Intrusion Sets are sets of related Incidents that are given an increase in attention and resources to detect, understand and respond [12]. Visualization for Correlation Analysis requires Timeline view for "...search...in current and historical data…" [12] and Flexibility for "...searches for patterns and trends..." [18]. Timeline displays an order of incidents that have taken placed over a period of time [11], used to coordinate all views [3] and Flexibility of visualization gives the ability to manipulate the focus point [11] to support the analytical process [1]. Another important feature for Correlation Analysis is having capabilities for Investigation for "...retrospectively reviewing...data...looking for unexplained patterns" [12]. Investigation capabilities would allow users to investigate data by supporting simultaneous investigations [1] by providing extensive capabilities for vulnerability assessment [11] and a platform for visually clarified distinctions between vulnerabilities and alerts [3]. This would lead to an effective visualization for a Site-Specific Analyst or a Tactical Defender performing Correlation Analysis.

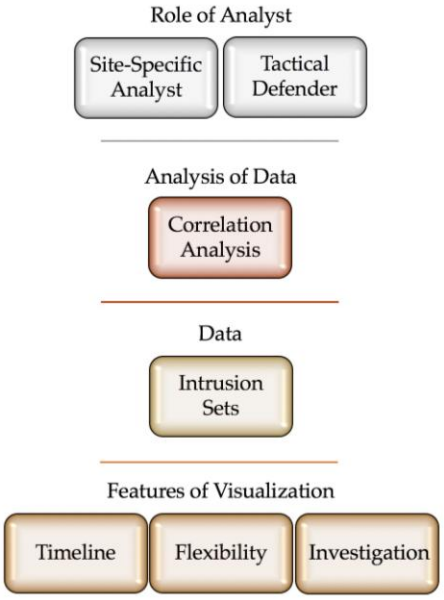

Figure 5. Visual representation of cognitive relationship for Correlation Analysis, as derived from the framework

5.4 Guidelines for effective visualization of threat analysis

Threat Analysis is an intelligent analysis [12] using additional data sources to profile attackers and their motivations [18].

A visual representation of this relationship is displayed in Figure 6. The excerpts and codes that led to this relationship are explained as follows: Threat Analysis is performed by Threat Analyst [12] along with Tactical Defender and Strategic Analyst [1]. A Strategic Analyst works at the community level [12] to understand the implications of an attack and categorize it [1]. It also works with Intrusion Sets as a type of Data as "Once incidents are confirmed...moves to...threat analysis..." [12]. Visualization for Threat Analysis requires Correlation as it uses "...additional data sources..." [18] and Interoperation as it uses "[additional other tools]...to gain additional insight..." [12]. Correlation visualizes relationships between different data dimensions to improve analyst performance [1]. This would lead to an effective visualization for a Threat Analyst, a Tactical Defender or a Strategic Analyst performing Threat Analysis.

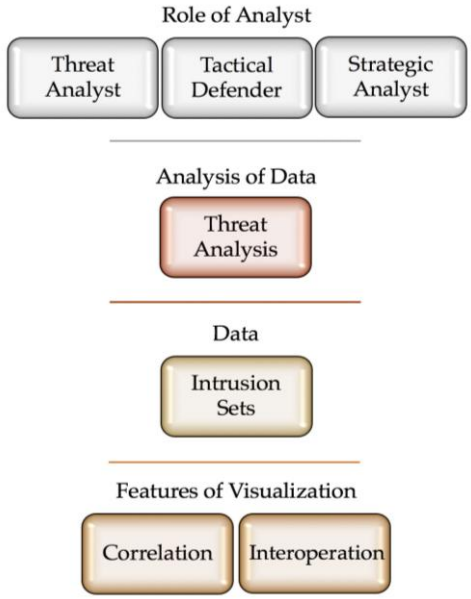

Figure 6. Visual representation of cognitive relationship for Threat Analysis, as derived from the framework

5.5 Guidelines for effective visualization of incident response analysis

Incident Response Analysis is when the analyst recommends or implements actions against a confirmed incident [12, 18].

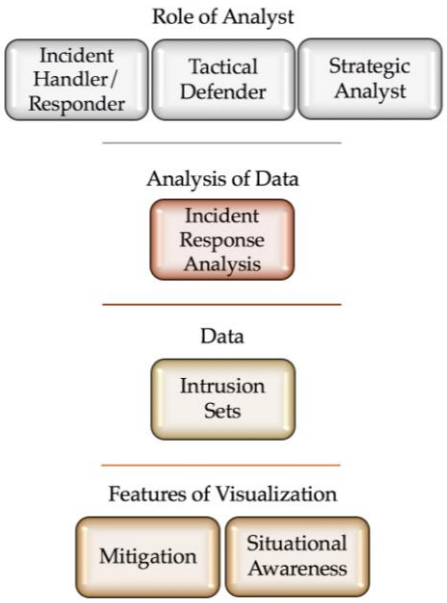

Figure 7. Visual representation of cognitive relationship for Incident Response Analysis, as derived from the framework

A visual representation of this relationship is displayed in Figure 7. The excerpts and codes that led to this relationship are explained as follows: Incident Response Analysis is usually performed by Incident Handler/Responder [12] along with Tactical Defender or Strategic Analyst [1]. It is a "...reaction to a confirmed incident" [18] and hence uses Intrusion Sets as a type of Data. Visualization for Incident Response Analysis requires Mitigation as it "...recommends and/or implements a course of action..." [12]. Mitigation capabilities would perform clean-up and containment; also, provide mitigation solution and/or activities [11]. Situational Awareness, is another feature of visualization, as it "...involves assessing the tradeoffs of potential responses and how the responses will impact organizational mission" [12]. This would lead to an effective visualization for an Incident Handler/ Responder, a Tactical Defender or a Strategic Analyst performing Incident Response Analysis.

5.6 Guidelines for effective visualization of forensic analysis

Forensic Analysis is gathering and preservation of data to support law enforcement agencies [12, 18]. It may take from hours to a few weeks to perform [11].

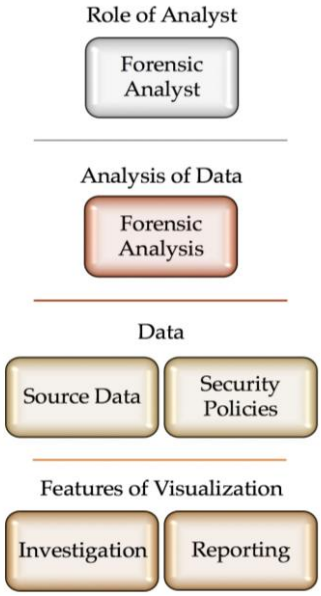

Figure 8. Visual representation of cognitive relationship for Forensic Analysis, as derived from the framework

A visual representation of this relationship is displayed in Figure 8. The excerpts and codes that led to this relationship are explained as follows: Forensic Analysis is performed by Forensic Analyst [12]. It "...preserves evidence in support of a law enforcement investigation" [18] and hence uses Security Policies as type of Data. These are policies defined by the government or organizations [11]

relating to cyber security; also, includes cyber law. It would also require Source Data of the confirmed incident found. Visualization for Forensic Analysis requires Investigation for "...gathering evidence..." [12] and Reporting to create reports for law-enforcement agencies. This would lead to an effective visualization for a Forensic Analyst performing Forensic Analysis.

5.7 Guidelines for effective visualization for impact assessment

Impact Assessment is the task of identification of impact, damage and potential critical nodes that may be reachable after a breach [11].

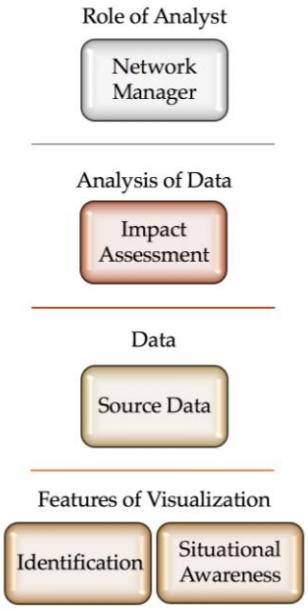

Figure 9. Visual representation of cognitive relationship for Impact Assessment, as derived from the framework

A visual representation of this relationship is displayed in Figure 9. The excerpts and codes that led to this relationship are explained as follows: Any analyst or Network Manager to perform Impact Assessment would use Source Data of the confirmed incident, as type of Data. For visualizing Impact Assessment, it would require Identification for "Impact identification... [identification of] mission impact and system impact..." [11]. It refers to the capabilities to identify vulnerabilities; malicious users; intended target of attacks; main resources of the system affected [11]. Situational Awareness, is also required, to find which "...domain is not protected enough..." [11]. This would lead to an effective visualization for a Network Manager performing Impact Assessment.

5.8 Guidelines for effective visualization for security quality management

Security Quality Management is the task related to services that support information security [18] in an organization like tutorials or training [11].

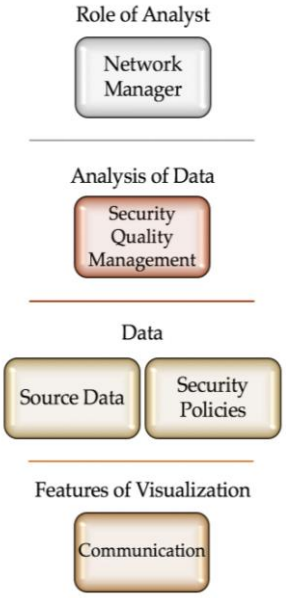

Figure 10. Visual representation of cognitive relationship for Security Quality Management, as derived from the framework

A visual representation of this relationship is displayed in Figure 10. The excerpts and codes that led to this relationship are explained as follows: Any analyst or Network Manager to perform Security Quality Management activities would use Source Data of the incident and Security Policies of the organization as types of Data. Visualizing Security Quality Management would require Communication as it includes "...services that support information security..." [12] and these [services] need to be communicated back to the Network Manager.

6. Conclusion

The literature review draws attention to a major issue in the field of cyber-security visualization vis-à-vis the lack of standardized evaluation techniques for effective visualization. The current evaluation techniques are neither standardized nor inclusive of user-validation. This leads to a cloud of uncertainty regarding the effectiveness of visualizations for cyber-security. Hence, there arose the need for a framework which appreciates the requirements of end users (security analysts) and evaluates the

effectiveness of cyber-security visualization for the performed task.

EEVi was developed to bridge the research gap by standardizing evaluation techniques using guidelines for effective cyber-security visualization for the performed task. These guidelines are formed as a result of the cognitive relationships or *links* associated with the performed task in the logic sequence derived from the structure of the framework (see Figure 2).

Using thematic analysis to develop EEVi led to the identification of storylines which represent guidelines for each task for effective visualizations. The guidelines for eight tasks were represented in this article. These eight tasks have been identified during the process of qualitative coding as these were the most common tasks conducted by security analysts.

The guidelines for effective visualization of these tasks have been confirmed by security analysts to define which features of visualization are effective to perform the task at hand, as a part of an expert review. These guidelines can thus be used to as a common framework to evaluate the effectiveness of visualizations developed for cyber-security.

Visualization developers can use the guidelines as fundamentals for developing effective cyber-security visualizations and use these guidelines to steer them in the direction of creating visualizations that would be effective for security analysts to perform their tasks. The guidelines can also form a basis of dialogue between the developer and end user (security analyst) to understand the requirements of security analysts in a manner that is understood by both sides.

Therefore, EEVi presents an effective solution to the research gap of lack of guidelines and a common framework to standardize evaluation for visualization of cyber security, and aims to increase adoption rates of cyber-security visualization by security analysts.

8. References

[1] G. Fink, C. North, A. Endert and S. Rose (2009) 'Visualizing cyber security: Usable workspaces', in Proceedings of the 6th International Workshop on Visualization for Cyber Security, IEEE, pp. 45–56.

[2] D. Staheli, T. Yu, R. J. Crouser, S. Damodaran, K. Nam, D. O'Gwynn, S. McKenna and L. Harrison (2014) 'Visualization evaluation for cyber security: Trends and future directions', in Proceedings of the Eleventh Workshop on Visualization for Cyber Security, ACM, pp. 49-56.

[3] S. Mckenna, D. Staheli and M. Meyer (2015) 'Unlocking user-centered design methods for building cyber security visualizations', in Proceedings of 2015 IEEE Symposium on Visualization for Cyber Security, IEEE, pp. 1–8.

[4] D. M. Best, A. Endert and D. Kidwell (2014) '7 key challenges for visualization in cyber network defense' in Proceedings of the Eleventh Workshop on Visualization for Cyber Security, ACM, pp. 33–40.

[5] S. Creese, M. Goldsmith, N. Moffat, J. Happa and I. Agrafiotis (2013) 'CyberVis: Visualizing the potential impact of cyber attacks on the wider enterprise', in Proceedings of the 2013 IEEE Conference on Technologies for Homeland Security, IEEE, pp. 73-79.

[6] M. Angelini, N. Prigent and G. Santucci (2015) 'Percival: proactive and reactive attack and response assessment for cyber incidents using visual analytics', in Proceedings of 2015 IEEE Symposium on Visualization for Cyber Security, IEEE, pp. 1–8.

[7] F. Roveta, G. Caviglia, L. Di Mario, S. Zanero, F. Maggi and P. Ciuccarelli (2011) 'Burn: Baring unknown rogue networks', in Proceedings of the 8th International Symposium on Visualization for Cyber Security, ACM, pp. 6:1–6:10.

[8] T. Wüchner, A. Pretschner and M. Ochoa (2014) 'Davast: Data-centric system level activity visualization', in Proceedings of the Eleventh Workshop on Visualization for Cyber Security, ACM, pp. 25–32.

[9] J. B. Colombe and G. Stephens (2004) 'Statistical profiling and visualization for detection of malicious insider attacks on computer networks', in Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, ACM, pp. 138–142.

[10] I. Agrafiotis, J. R. Nurse, O. Buckley, P. Legg, S. Creese and M. Goldsmith (2015) 'Identifying attack patterns for insider threat detection', Computer Fraud & Security, Elsevier 2015(7), July 2015, pp. 9–17.

[11] R. F. Erbacher, D. A. Frincke, P. C. Wong, S. Moody and G. Fink (2010) 'A multi-phase network situational awareness cognitive task analysis', Information Visualization, ACM 9(3), June 2010, pp. 204–219.

[12] A. D'Amico and K. Whitley (2008) 'The Real Work of Computer Network Defense Analysts', in VizSEC 2007: Proceedings of the Workshop on Visualization for Computer Security, Springer, pp. 19-37.

[13] H. Lam, E. Bertini, P. Isenberg, C. Plaisant and S. Carpendale (2012) 'Empirical Studies in Information Visualization: Seven Scenarios,' IEEE Transactions on Visualization and Computer Graphics, IEEE 18(9), Sept. 2012, pp. 1520-1536.

[14] I. Vessey (2015) 'The theory of cognitive fit: One aspect of a general theory of problem-solving', in Human-

computer Interaction and Management Information Systems: Foundations, Chapter 8, Taylor & Francis, pp. 141–183.

[15] V. Braun and V. Clarke (2006) 'Using thematic analysis in psychology', Qualitative Research in Psychology, Edward Arnold (Publishers) Ltd 3(2), pp. 77–101.

[16] F. M. Albar and A. J. Jetter (2013) 'Uncovering project screening heuristics with cognitive task analysis: How do gatekeepers decide which technologies to promote?', in 2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services, IEEE, pp. 459–467.

[17] B. Crandall, G. Klein, and R. Hoffman (2006) 'Working Minds: A Practitioner's Guide to Cognitive Task Analysis', MIT Press.

[18] A. D'Amico, K. Whitley, D. Tesone, B. O'Brien and E. Roth (2005) 'Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts', Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE 49(3), pp. 229–233.

[19] M. Vaismoradi, J. Jones, H. Turunen and S. Snelgrove (2016) 'Theme development in qualitative content analysis and thematic analysis', Journal of Nursing Education and Practice, Sciedu Press 6(5), pp. 100-110.

9. Acknowledgements

We thank the University of Southampton for sponsoring and providing the necessary resources and guidance to perform this research. We would also like to give a special thanks to Giles Howard, Toby Wilkinson and Nawfal Fadhel for proofreading our work and giving useful feedback.

