2018

Novel Alert Visualization: The Development of a Visual Analytics Prototype for Mitigation of Malicious Insider Cyber Threats

Karla A. Clarke
Nova Southeastern University, kclarke888@gmail.com

Novel Alert Visualization: The Development of a Visual Analytics Prototype for Mitigation of Malicious Insider Cyber Threats

by

Karla A. Clarke

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Information Systems

College of Engineering and Computing
Nova Southeastern University

2018

We hereby certify that this dissertation, submitted by Karla Clarke, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.

Yair Levy, Ph.D.
Chairperson of Dissertation Committee

6/4/2018
Date

Shonda Brown, Ph.D.
Dissertation Committee Member

6/4/2018
Date

Laurie P. Dringus, Ph.D.
Dissertation Committee Member

6/4/2018
Date

Approved:

Yong X. Tao, Ph.D., P.E., FASME
Dean, College of Engineering and Computing

6/4/2018
Date

College of Engineering and Computing
Nova Southeastern University

2018

An Abstract of a Dissertation Submitted to Nova Southeastern University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Novel Alert Visualization: The Development of a Visual Analytics Prototype for Mitigation of Malicious Insider Cyber Threats

by
Karla A. Clarke

June 2018

Cyber insider threat is one of the most difficult risks to mitigate in organizations. However, innovative validated visualizations for cyber analysts to better decipher and react to detected anomalies has not been reported in literature or in industry. Attacks caused by malicious insiders can cause millions of dollars in losses to an organization. Though there have been advances in Intrusion Detection Systems (IDSs) over the last three decades, traditional IDSs do not specialize in anomaly identification caused by insiders. There is also a profuse amount of data being presented to cyber analysts when deciphering big data and reacting to data breach incidents using complex information systems.

Information visualization is pertinent to the identification and mitigation of malicious cyber insider threats. The main goal of this study was to develop and validate, using Subject Matter Experts (SME), an executive insider threat dashboard visualization prototype. Using the developed prototype, an experimental study was conducted, which aimed to assess the perceived effectiveness in enhancing the analysts' interface when complex data correlations are presented to mitigate malicious insiders cyber threats.

Dashboard-based visualization techniques could be used to give full visibility of network progress and problems in real-time, especially within complex and stressful environments. For instance, in an Emergency Room (ER), there are four main vital signs used for urgent patient triage. Cybersecurity vital signs can give cyber analysts clear focal points during high severity issues. Pilots must expeditiously reference the Heads Up Display (HUD), which presents only key indicators to make critical decisions during unwarranted deviations or an immediate threat.

Current dashboard-based visualization techniques have yet to be fully validated within the field of cybersecurity. This study developed a visualization prototype based on SME input utilizing the Delphi method. SMEs validated the perceived effectiveness of several different types of the developed visualization dashboard. Quantitative analysis of SME's perceived effectiveness via self-reported value and satisfaction data as well as qualitative analysis of feedback provided during the experiments using the prototype developed were performed.

This study identified critical cyber visualization variables and identified visualization techniques. The identifications were then used to develop QUICK.v™ a prototype to be used when mitigating potentially malicious cyber insider threats. The perceived effectiveness of QUICK.v™ was then validated. Insights from this study can aid organizations in enhancing cybersecurity dashboard visualizations by depicting only critical cybersecurity vital signs.

Table of Contents

List of Tables

Tables

List of Figures

Figures

Chapter 1

Introduction

Background

Big data analytics is altering cybersecurity in a potentially disruptive way by introducing profuse amounts of massively incomplete sets of data (Kott, Swami, & McDaniel, 2014). Big data is also used to detect the threat of a cyber attack or potentially cyberterrorism, hence, the need for data to be protected from unauthorized access, use, or manipulation (AlMutairi, Abdullah, AlBukhary, & Kar, 2015). Data is accelerating remarkably fast (Geer Jr., 2011). Data is defined as "a subset of information in electronic format that may be retrieved or transmitted" (NIST, 2013). "As information spaces expand in size and complexity, there is a growing need for visual representations that help us make better sense of diverse data relations and patterns" (Dork, Carpendale, & Williamson, 2011, p. 20). There are many challenges faced when deciphering large volumes and varieties of data within complex information systems (Shneiderman & Plaisant, 2015). A particularly complex challenge faced is dealing with malicious insider cyber threats (Pfleeger & Stolfo, 2009). Malicious cyber insiders may aim to inconspicuously exfiltrate large volumes of data within an organization (Agrafiotis et al., 2015). Thus, cyber analysts are presented profuse amounts of alerts when using data visualizations to address the challenge of deciphering and reacting to big data within complex information systems, resulting in information overload ("Big Data Meets", 2012).

Quickly analyzing overwhelming amounts of data and responding during a malicious insider attack not only requires analytics tools but also human judgment (Gorg, Kang, Liu, & Stasko, 2013). According to Gorg et al. (2013), "the analysis process requires human judgment to make the best possible evaluation of incomplete, inconsistent, and potentially deceptive information in the face of rapidly changing situations" (p. 30). Cyber analysts have to perceptively determine how to react during an attack, while they may be inadequately equipped based on the visualization being used. Analytics tools are most useful if the cyber analyst can focus their attention by utilizing a systematic and focused visualization, which aided in sharpening the analytic focus, essentially allowing analysts to find patterns and anomalies of interest (Shneiderman & Plaisant, 2015). Staheli, Mancuso, Leahy, and Kalke (2016) addressed data visualization challenges faced within the Department of Defense (DoD) by developing a cyber dashboard. They noted that within cybersecurity incorporating hundreds of data sources could be very difficult, requiring the need for this level of incorporation is a vital element for future designs. Future designs need to address the prevalent visual analytics challenge when examining data from multiple sources (Shneiderman & Plaisant, 2015). This study aims to develop and validate a visualization prototype that intends to enhance the presentation of complex data correlations. The study addressed the need for further analysis of end-user specifications for the development of a cybersecurity visualization dashboard (Inibhunu et al., 2016; Agrafiotis et al., 2015; McKenna, Staheli, & Meter, 2015). The findings of this research add to the Information Systems (IS) and Information Security (InfoSec) body of knowledge by developing a novel and effective detection method for the identification of anomalous activities when mitigating malicious insider

cyber threats.

A visualization dashboard was also developed that presents cybersecurity 'vital signs', by assessing the perceived effectiveness of enhancing the presentation of complex data correlations when mitigating malicious insiders cyber threats. Using SMEs, the cyber insider threat dashboard visualization prototype was developed and validated. The remainder of this dissertation is organized as follows. First, the research problem was presented. Followed by the main dissertation goal, research questions, relevance, and significance. Then, the barriers and issues as well as the definition of terms are presented. Next, the literature review is presented, followed by the methodology. Next, the results, conclusions, implications, recommendations, and summary are presented. Lastly, the appendices are presented.

Problem Statement

The research problem that this study addressed was the prevalent challenge faced within the cybersecurity industry when detecting potentially malicious insider cyber threats, to enable visualization of those threats as they occur (Gorg et al., 2013; Inibhunu et al., 2016; Pfleeger & Stolfo, 2009; Patcha & Park, 2007). The nature of insider threats remains unchanged within cybersecurity research, and it remains a complicated threat to mitigate (Kumarmandal & Chatterjee, 2015). "Employees and contractors are the second greatest threat to an organization, exceeded only by hackers", as such employees and contractors are considered insiders (Greitzer, Moore, Cappelli, Andrews, Carroll, & Hull, 2008, p. 61). A hacker is an unauthorized outsider who initiates a threat or attack (Sun, Srivastava, & Mock, 2006). An 'insider' has legitimate access to an organization

(Pfleeger & Stolfo, 2009). 'Insider threat' refers to, individuals with legitimate access whose behaviors put data, intellectual property, systems, organizations, and businesses at risk of being compromised (Pfleeger, Predd, Hunker, & Bulford, 2010; Predd, Pfleeger, Hunker, & Bulford, 2008). A malicious insider "is an insider who has malicious intent that acts against the best interests of the organization" (Santos et al., 2012, p. 331).

Malicious insiders within cyberspace are significant challenges that organizations face (Azaria, Richardson, Kraus, & Subrahmanian, 2014). In 2005, a United States (U.S.) Justice Department survey found that 74% of all cyber-theft within organizations were carried out by insiders and 40% of all cyber-incidents reported by 36,000 U.S. businesses involved insiders (Rantala, 2008). In 2013, there were over 117,000 cyber attacks per day costing firms over $28 million (Price Waterhouse Cooper, 2014). Current insider threat detection solutions inevitably trigger large volumes of false positive alerts. A false positive alert is a false alarm triggered when a detected vulnerability does not actually exist but is counted in a measurement as valid, requiring investigation of the incident and organizational resources (Mell, Bergeron, & Henning, 2005). According to Victor, Rao, and Venkaiah (2010), most intrusion detection systems (IDSs) have very high rates of false positives. This may lead to desensitized analysts ignoring possibly dangerous exploits that pose potentially detrimental financial and intellectual property damage to organizations (Spathoulas & Katsikas, 2010). An intrusion may be caused by "attackers accessing systems from the Internet or by authorized users of systems that attempt to misuse the privileges given to them and/or to gain more privileges for which they are not authorized" (Lazarevic, Kumar, & Srivastava, 2005, p. 21). An IDS, defined as a misuse or anomaly detector, may be used to detect unauthorized or malicious attacks over a

system that primarily occurs through the Internet (Kumar & NandaMohan, 2008).

Traditional intrusion detection and prevention systems may be insufficiently designed -- they may not be capable of promptly identifying malicious insiders cyber threats or they generate a considerable amount of false positive alerts (Agrafiotis et al., 2015; Spathoulas & Katsikas, 2010). Since intrusive activity does not always correlate with anomalous activity, newly developed insider threat solutions that use the same techniques as traditional IDSs may not be adequate. Newly developed insider threat solutions should allow for rapid analysis of complex data correlations relevant to the identification of potentially malicious cyber insiders (Patcha & Park, 2007). While IDSs are good at detecting intrusions they do not specialize in anomaly identification (Gandomi & Haider, 2015). Anomaly detection can be very difficult within large complex data sets as it may result in spurious correlations or "uncorrelated variables being falsely found to be correlated due to the massive size of the dataset" (Gandomi & Haider, 2015, p. 143). An anomaly detection system aids with the identification of abnormal behaviors based on complex data correlations. In this study a complex data correlation refers to identifying linear or non-linear relationships between two or more data variables (Patcha & Park, 2007).

Detecting malicious cyber insider threats is a complex task since their malicious actions take place along normal activities (Azaria et al., 2014). Identifying anomalous activities amidst appropriate activities pose the potential difficulty of being able to identify legitimate anomalies within the data presented. Shneiderman and Plaisant (2015) referred to this difficulty as the 'analytic-focusing problem'. Useful visualizations should sharpen the analytic focus to enable detection of patterns or anomalies of interest

(Shneiderman & Plaisant, 2015). Detecting misuse by malicious cyber insiders involves examining an individual's use of information and resources to decide whether the use is legitimate and/or deviates from what is considered 'normal' activities (Caputo, Maloof, & Stephens, 2009). Cyber analysts must find outliers or anomalies (also referred to as anomalous activities) within all of the users generated activities utilizing data analytics and information visualization (Kang, Gorg, & Stasko, 2011). Data analytics also uses data mining of large volumes of records, images, and activities translated to emphasize areas of interest that aid in understanding complex data (Kang et al., 2011).

Researchers are combining data mining and information-visualization to allow for visual inspection of data examination of outlier data (Shneiderman, Plaisant, Cohen, & Jacobs, 2010, p. 556). Information visualization is "communicating and perceiving data, both abstract and scientific, through visual representations" (Roberts et al., 2014, p. 27). Visual analytics supports human decision-making in complex application domains (Arias-Hernández, Dill, Fisher, & Green, 2011). For instance, visual analytics is used in medicine for anomaly detection within patients' vital signs (Dutta, Maeder, & Basilakis, 2013). There are challenges faced with "archival, retrieval, and transformation" of the data when deciphering data analytics logs (Levy & Ramim, 2012, p. 99). Misspelled, duplicated, or data written in foreign languages may enhance the difficulty of misuse detection within an abundance of data (Jonas, 2006). Imbalanced data also often results in very high accuracy for the majority class (benign users), and very low accuracy for the minority class (malicious insiders). Identifying anomalies while taking corrective action accordingly poses a challenge (Azaria et al., 2014).

Cyber analysts have to prioritize triggered alerts related to data moved over the

Internet and account for relevant cyber attacks. Security Information and Event Management (SIEM) tools are applications that offer the ability to gather security data from information system components and present that data as actionable information via a single interface. SIEM tools are used by many cyber analysts within large organizations (NIST, 2013). Development techniques for visualization systems in the context of cybersecurity tools should be evaluated since increased data is driving endless alerts most are falsely noted and increasing the response time required for decision-making (Arias-Hernández et al., 2011). In a case study performed by ACI Payment Systems (2015), when the Canadian Federal Government enforced stricter anti-money laundering regulations. By applying enhanced profiling tables and alert rules to an existing risk management solution reduced debit card fraud alerts by 84% and reduced analyst resources by 50%. These findings indicate an enhancement of profiling tables produced less work for analysts while increasing detection rates. There is a lack of effective methods for analysts to investigate events generated from big data infrastructure equipment and to find the correct diagnosis of critical alert information from the excessive alerts (Boukri & Chaoui, 2015).

Prior research that examined detection of malicious insider cyber threats has mainly been focused on anomaly detection methods, malicious behaviors, or detection techniques used by cyber analysts (Agrafiotis et al., 2015; Azaria et al., 2014; Legg et al., 2015; Santos et al., 2012). Kemmerer and Vigna (2002) defined anomaly detection as, "models of the intended users and applications behaviors, which interpret deviations from normal behavior as a problem" (p. 28). According to Legg et al. (2015), cyber analysts can be empowered to identify anomalous activities by combining detection results with a

visual analytics approach for expedited real-time detection. Arias-Hernández et al. (2011) concluded that visual analytics challenges would require coordination of Human Computer Interaction (HCI) practitioners as well as visualization and computation researchers. Boukri and Chaoui (2015) also conferred that dashboard-based visualization techniques could be used to provide full visibility of network progress and problems in real-time. Heckman, Stech, Schmoker, and Thomas (2015) speculated that conventional approaches to cybersecurity are inadequate since conventional approaches that may include applying firewalls and IDS technologies are still being penetrated allowing sensitive information to be exploited. Accordingly, there is an evident need for an unconventional approach to address cybersecurity challenges when detecting cyber insider threats. This study approached the challenge of detecting cyber insider threats in an unconventional way by utilizing a newly developed visualization prototype to enhance the presentation of complex malicious insider cyber threat, both linear and non-linear, indicator correlations.

In the field of medicine, practitioners understand the importance of monitoring as well as recording patients' vital signs (Mok, Wang, & Liaw, 2015). In most cases just four vital signs serve as an essential tool for determining life saving responses (Harries, Zachariah, Kapur, Jahn, & Enarson, 2009). Within business organizations, Executive Dashboards (EDs) are used to present real-time information pertinent to the organizations strategy and risks because EDs are designed to enable efficient decision-making (Ballou, Heitger, & Donnell, 2010). Within the Department of Defense (DoD) the Under Secretary of Defense Comptroller's office implemented EDs for presenting credible and timely data where only core data with concise tables are presented to analysts (Dees,

2009). Similarly, the Heads Up Display (HUD) considerably altered cockpit information for pilots. Using cross-monitoring principles a pilot continually references external visual cues, correlations of pitch, attitude, and vertical speed to make decisions in the event of unwarranted deviations (Dopping-Hepenstal, 1981). This is pertinent to highlighting how visualizations are already being used within these well-established industries. Practitioners are making critical decisions daily using visualizations and EDs. Shneiderman et al. (2010) indicated, "new visualization products should be more than just cool, they should offer measurable benefits for realistic tasks" (p. 539). There is an apparent need to identify and validate requirements and define visualization techniques for the development of a cyber insider threat visualization prototype (Legg et al., 2015; Shneiderman & Plaisant, 2015). As EDs, HUDs, and vital signs have enhanced their industries, the cyber insider threat visualization prototype developed in this study similarly aims to enhance the visualizations presented in the cybersecurity industry.

Dissertation Goal

The main goal of this study was to develop and validate, using Subject Matter Experts (SMEs), a cyber insider threat dashboard visualization prototype. The prototype used in an experimental study that aimed to assess the perceived effectiveness of enhancing the presentation of complex data correlations when mitigating malicious insiders cyber threats. The need for this work is demonstrated by the research of Albanese, Pugliese, and Subrahmanian (2013), Boukri and Chaoui (2015), Dork et al. (2011), Greitzer and Hohimer (2011), Legg et al. (2015), Shneiderman et al. (2010), as well as Shneiderman and Plaisant (2015). Albanese et al. (2013) developed a graphical

index that would aid in providing evidence of occurrences of an activity, and identify if an anomaly matches a sequence of observations in a simulated environment. Dork et al. (2011) approached the challenge of complex data visualization by presenting a novel visualization technique combining implicit and explicit data relationships and still finding that more work is necessary to understand complex information spaces. Though the threat posed by malicious cyber insiders is real, there is a lack of precise analysis due to the sheer volume of activities (Legg et al., 2015; Spathoulas & Katsikas, 2010).

Prior research that has developed dashboard visualization prototypes rarely addressed the perceived effectiveness using multiple SMEs within the applicable field (Goodall, 2007). Without a deeper understanding of the cyber analyst using the visualization, progress and practical application was difficult (Gorg et al., 2013). Assessing the perceived effectiveness of the developed cyber insider threat dashboard visualization prototype enhanced the validity for this research (Petter, Delone, & McLean, 2013). The perceived effectiveness was determined by obtaining a rating for satisfaction and value of the developed prototype (Hong, Tai, Hwang, Kuo, & Chen, 2017; Levy, 2006; Ellis & Levy, 2009). Identifying and assessing the risk score was beyond the scope of this study. Risk calculation are addressed in the AI-InCyThR™ prototype (Hueca, Clarke, & Levy, 2016). Once the prototype was developed, cybersecurity analysts were asked to assess the perceived effectiveness based on their user experience. The goal of the developed prototype is to alleviate the issues faced when using visualizations to identify potentially malicious cyber insiders.

There are multiple issues that reside with IDSs, for instance, triggering alarms where over 90% of the alerts are false positives, while identifying the true positive alerts

can be error prone and labor intensive (Ho, Lai, Chen, Wang, & Tai, 2012). This makes it difficult for cyber analysts to identify alerts that are legitimate, important, and procure the appropriate response (Julisch, 2003). Another issue faced is identifying anomalies within the environment (Patcha & Park, 2007). Visualizing the data can aid in presenting strong connections between individuals and entities, as was utilized to explore the Commission Report about the September 11, 2001 terrorist attacks (Gorg et al., 2013). Applying visualizations may assist with the issues encountered with current IDSs. "Visualizations use enormous visual bandwidth and a remarkable human perceptual system to enable users to make discoveries, decisions, or propose explanations" (Shneiderman et al., 2010, p. 538). Legg et al. (2013) developed a conceptual model for insider threat that would give an all-encompassing organizational view. Applying the proper visualizations within cybersecurity may aid management decision-making efforts. It appears there is a need for a visualization prototype that presents actionable alerts to cyber analysts.

Greitzer and Hohimer (2011) found that in a substantial number of cases, prior to an exploit, the malicious intent of the insider was 'observable'. Legg et al. (2015) later suggested that visualizations enabled cyber analysts to identify what particular attributes caused the insider to be detected, and coupling the detection results with visual analytics enhanced detection of anomalous activities. Pfleeger et al. (2010) concluded that by using four basic dimensions of the insider threat (the organization, the system, the individual, and the environment) insider actions could be evaluated to frame responses. In Figure 1, these four basic dimensions are used to frame the Quality User Insider ChecKing visualization (QUICK.v™) interface that presented identified vital signs pertinent to the identification of malicious cyber insiders. Applying the concept of four essential vital

signs from the medical field, the primary visualization technique for each identified cyber vital sign was selected and applied to QUICK.v™. Based on Shneiderman (1996)'s visualization mantra: overview, zoom and filter, then details on demand. Figure 2 presents a Sankey visualization that allows for zoom, filter of generated details, and may also allow analysts access to complete details on demand. Shneiderman et al. (2010) noted that new visualization products should "present information more rapidly and allow user-controlled exploration", without overwhelming novice users (p. 558). This study builds on earlier research by Boukri and Chaoui (2015), as well as Pfleeger et al. (2010) by developing and testing a visual analytics prototype that aims to decrease the time it takes to react to potentially malicious cyber activities.

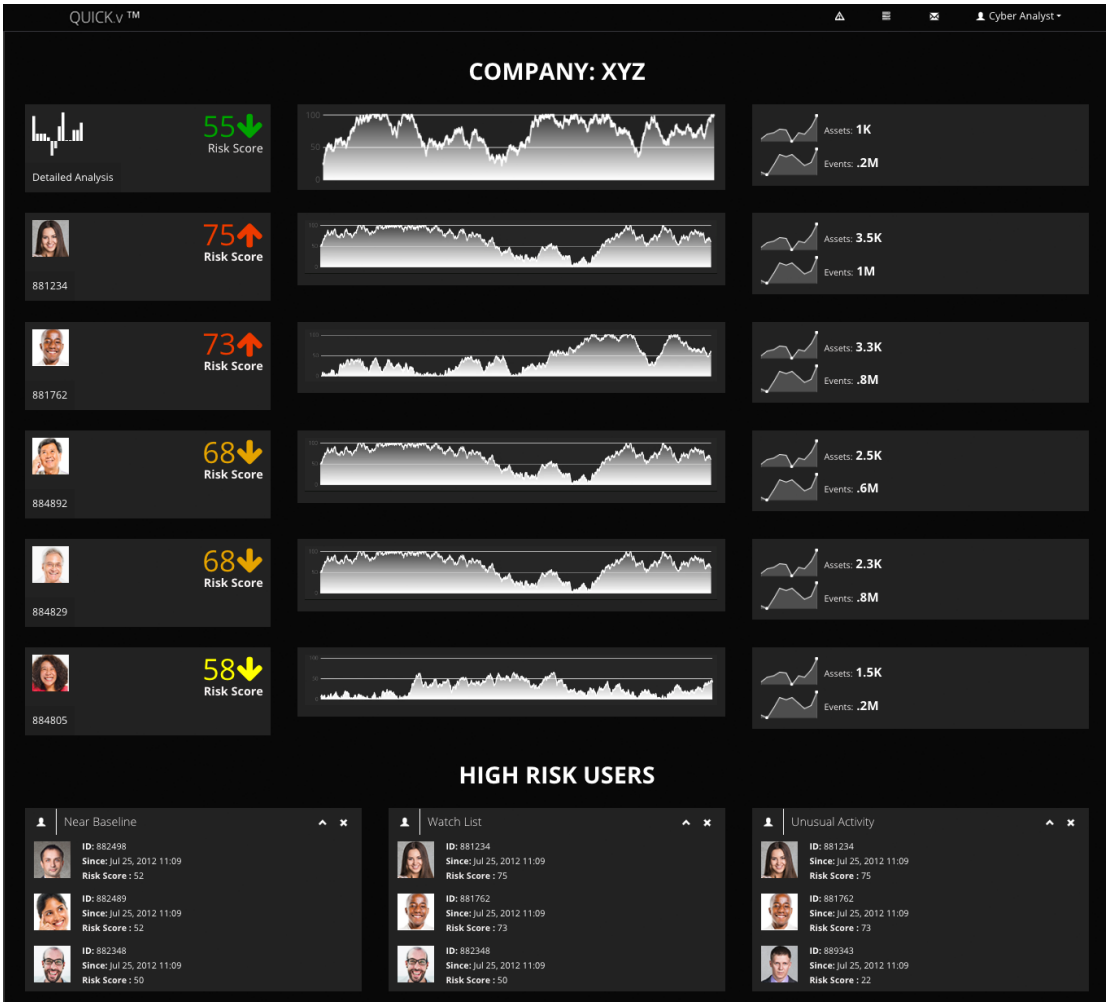


Figure 1: Quality User Insider ChecKing visualization (QUICK.v™) interface Dashboard

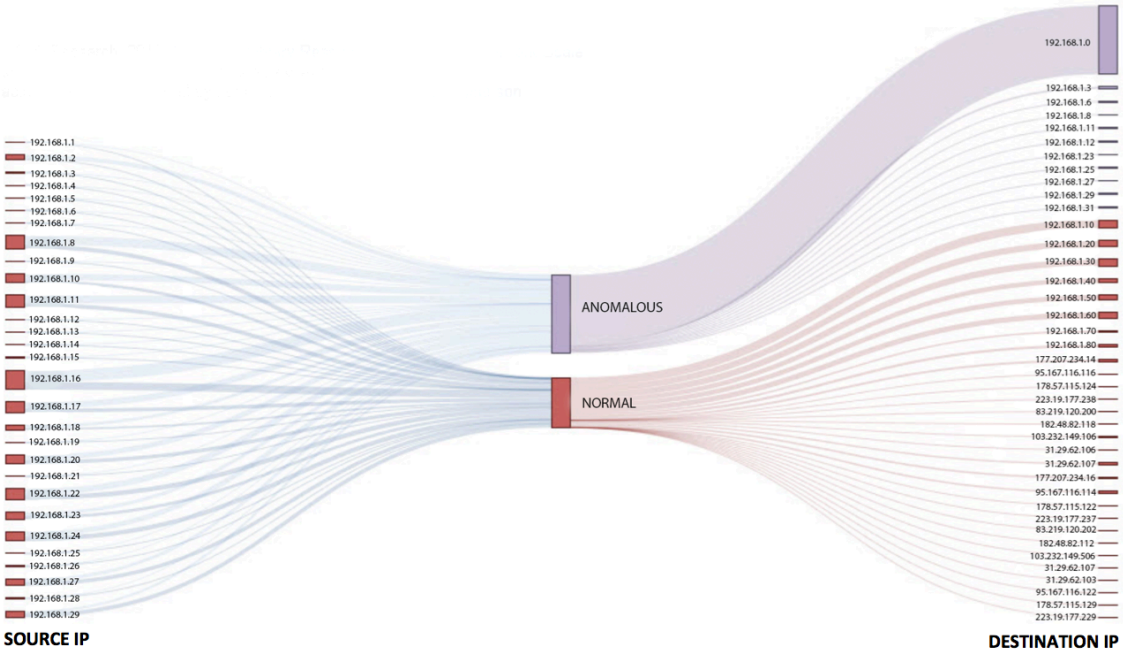


Figure 2: QUICK.v™ Interface Detailed Analysis

This study had five specific goals. The first research goal was to identify, using SMEs, the critical cyber visualization variables (see Appendix A). The cyber visualization variables refer to analytic variables that may aid in identifying potentially malicious cyber insiders (Casey, 2015). The second research goal was to identify, using SMEs, the rank order of the critical cyber visualization variables that the developed prototype should include, which may aid in identifying potentially malicious cyber insiders. The third research goal was to identify, using SMEs, the most valid presentation of complex data correlations using the identified critical visualization variables over multiple visualization techniques. The fourth research goal was to apply SMEs' identified critical visualization variables, in rank order, and techniques to develop QUICK.v™. The fifth research goal was to conduct an experimental study using SMEs to assess the

perceived effectiveness using self-reported value and satisfaction of the QUICK.v™ prototype when mitigating malicious cyber insiders.

Research Questions

In the medical field, vital signs are the core component of clinical management, aiding healthcare professionals with identifying potential threats to the normal operations of the body (Harries et al., 2009). Tracking insiders within cyber and the ability to analyze normal insider activities or when insiders are posing a threat appears to be much needed (Legg et al., 2015). It is pertinent to identify changes in vital signs early as this may prevent patient deterioration (Mok et al., 2015). Likewise, identifying malicious cyber insiders early may prevent an organization from financial decline. The overarching research question that this study addressed is: What visualization variables and techniques should be used to develop a cyber insider threat dashboard visualization prototype that enhances the effective presentation of complex data correlations within cybersecurity? There is an abundance of visualization techniques available: graph-based, hierarchical, pixel-oriented, icon-based, geometric, etc. (Keim, 2000). Using the most effective visualization technique is vital for enhancing cyber insider threat analysis. The specific research questions that this study addressed are:

RQ1: What are SMEs' identified critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats?

RQ2: What is the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat

dashboard visualization prototype to detect potentially malicious cyber insider activities?

RQ3: What SMEs' identified visualization techniques are most valid to present complex cyber data *correlations* relevant to the pre-designated critical cyber visualization variables that are applied within the developed cyber visualization prototype QUICK.v™?

RQ4: What SMEs' identified visualization techniques are most valid to present *top six critical cyber visualization variables* to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™?

RQ5: What is the SMEs' perceived effectiveness (i.e. satisfaction & value/importance) of the QUICK.v™ prototype when mitigating potentially malicious cyber insider threats?

Utilizing a cyber visualization prototype like QUICK.v™ may allow cyber analysts to detect malicious insider threats more efficiently, while allowing cyber leaders to take effective corrective actions. By reducing the key indicators of malicious activities to be displayed specifically for cyber analysts, QUICK.v™ may enable early detection and prevention of malicious cyber insider threats. Therefore, a real-time overall view of the cyber environment using appropriate visualization techniques aided in addressing visual analytic challenges faced by cyber analysts (Boukri & Chaoui, 2015).

Relevance and Significance

Relevance

This study presents a novel way of addressing the prevailing problem faced when detecting malicious cyber insiders. Past studies have aimed at addressing this challenge, however, the problem persists. Randazzo, Keeney, and Kowalski (2005) detailed numerous insider attacks from 1996 to 2002, with losses to individual firms ranging from $200 thousand to $600 million. According to Pfleeger et al. (2010), "insider misuse can threaten personal data, national security, as well as economic prosperity" (p. 169). Recently, this challenge has still continued to persist. Price WaterHouse Coopers (PwC) identified that insiders are likely to be to the primary source of cyber attacks (PwC, 2013). In 2014, a significant percentage of U.S. executives worried that cyber threat would impact growth within their organizations (PwC, 2014). Approaching this problem from an alternate perspective with a novel development can aid in addressing this problem. Cole (2015) conducted a survey on 772 individuals based on the results, 34% of respondents estimated over $1 million was lost due to insider attacks. Therefore, identifying potential mitigations for malicious cyber insider threats is pertinent to preventing detrimental financial and intellectual property damage to organizations. By applying a novel approach to solving this problem, new insight into enhancing mitigations by utilizing appropriate visualization techniques may be applied in order to detect malicious cyber insiders.

Significance

This study assisted in the identification of malicious insiders (Agrafiotis et al., 2015; Pfleeger & Stolfo, 2009). This was done by obtaining user requirements from

current cyber SMEs, applying the requirements to the development of a dashboard visualization prototype, and iterating the design, using SMEs to determine the perceived effectiveness of the developed prototype. Similar to cyber analysts, nurses perform surveillance to identify changes in activity and protect patients from harm (Rogers, Dean, Hwang, & Scott, 2008). Additionally, communication delays of vital signs may result in patient deterioration (DeVita, 2005). Within cybersecurity mitigating adverse incidents require surveillance to identify anomaly metrics and attack patterns (Agrafiotis et al., 2015). Delays in identifying a potentially malicious cyber insider may result in substantial losses, resulting in the deterioration of an organization (Randazzo et al., 2005). Developing dashboard visualizations similar to those used by nurses when making critical decisions, but tailored toward cyber vitals incorporates a novel approach to anomaly detection within cybersecurity. Therefore, to develop a cyber dashboard prototype for the detection of malicious cyber insiders, input was obtained from cyber SMEs with the knowledge and experience to identify effective indicators, data points, and fluctuations allowing for the development of a more relevant and valid prototype. This study is substantially significant since there are substantial financial losses resulting from exploits perpetrated as a result of malicious cyber insiders.

Barriers and Issues

A barrier was collecting results from the SMEs over multiple iterations, to obtain the pertinent cyber variables and then to validate the developed prototype. As a result, participants were rewarded with a gift card for participation. Another barrier of this study was that after administering the Delphi technique within the second phase, an experiment

then needed to be conducted applying the developed prototype for SME assessment. Since the participants were all volunteers, they could have withdrawn from the study at any time skewing the final results (Ellis & Levy, 2009). By rewarding participants for their time this reduced the rate of withdrawal. Additionally, variations in the SMEs years of experience may alter the informality of their results, as some SMEs may be more experienced than others. As such some SMEs may be more familiar with the requirements necessary for the adequate detection and remediation of cyber insider threats than others. To ensure all SMEs had a baseline understanding of insider threats, prior to conducting the survey, all SMEs were informed of what insider threat means within the scope of this study, as well as identified variables for anomaly identification within cybersecurity, and significant benefits of mitigating malicious cyber insider threats.

Limitation and Delimitation

Limitation

A limitation of this study was that the developed visualization prototype intended to visualize complex correlations based on cybersecurity related data. The cybersecurity data needed to be fed to the developed visualization prototype from viable data sources. The parsed data feeds were then utilized for generating the visualizations on the validated front-end. If the data input "was either incorrect, of low quality, or irrelevant, the resulted output was going to be ineffective regardless of the quality of the processing, colloquially, garbage-in/garbage-out" (Levy & Ellis, 2006, p. 185). If the data source was corrupted or incorrect the visualizations presented would be inaccurate. The findings of

this study served as the foundation of what needs to be presented within current cybersecurity visualizations. The prototype developed served as a tool for simplifying current visualization techniques and presenting the generated events. However, as the source of the data feeds change overtime the visualization variables used to develop the prototype may need additional validation. The prototype being developed would represent variables relevant to current applications and data sources. Therefore, future research may be required to apply the prototype that was developed using SMEs to future data sources.

Delimitation

A potential delimitation of this study was that the developed prototype did not perform data parsing the developed prototype visualized already parsed data in a consumable form factor pertinent to cyber professionals. Additionally, this study was limited to participants who have worked or are working within cybersecurity.

Definition of Terms

The following represent terms and definitions.

Anomaly Detection – "Models of the intended users and applications behaviors that interpret deviations from normal behavior as a problem" (Kemmerer & Vigna, 2002, p. 28).

Cyber Attack – "An attack, via cyberspace, that targets an enterprise's use of cyberspace for the purpose of disrupting, destroying, or maliciously controlling a computer environment/infrastructure; destroying the integrity of the data; or stealing controlled information" (NIST, 2013, p. 57).

Cyber Crime – "Any crime that involves computers and networks, including crimes that do not rely heavily on computers" (Casey, 2000, p. 8).

Cybersecurity – "Prevention of damage to, unauthorized use of, exploitation of, and, if needed, the restoration of electronic information and communications systems to ensure confidentiality, integrity, and availability" (Axelrod, 2006, p. 1).

Cyberspace or 'Cyber' – Independent network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries (The White House, 2009).

Cyberterrorism – "Concerted, sophisticated attacks on networks" (as cited in Foltz, 2004, p. 154).

Data - A subset of information in an electronic format that allows it to be retrieved or transmitted (NIST, 2013, p. 58).

Data Analytics – The use of data mining of large volumes of records, images, and activities translated to highlight areas of interest to aid in understanding complex data (Leventhal, 2010).

Data Breach - An organization's unauthorized or unintentional exposure, disclosure, or loss of sensitive PI, such as social security numbers; financial information, such as credit card numbers; date of birth; or mother's maiden name (NIST, 2010).

False-positive Alerts – "When something (e.g., a vulnerability) does not actually exist but is counted in a measurement" (Mell, Bergeron, & Henning, 2005, p. 40).

Information System (IS) – The system that governs the information technology development, use, application, and influence on a business or corporation (Alvarez, 2002).

Information Visualization – "Communicating and perceiving data, both abstract and scientific, through visual representations" (Roberts et al., 2014, p. 27).

Insider – "Individuals who have legitimate access to an organization" (Pfleeger, Predd, Hunker, & Bulford, 2010, p. 169).

Insider Attack - The abuse of privileges or access by an insider that results in a breach, interruption or disregard of a law, rule, or policy (Goodall, 2007).

Insider Threat – "Individuals with legitimate access whose behaviors put a firm's data, intellectual property, systems, organizations, and businesses at risk of being attacked" (Pfleeger, Predd, Hunker, & Bulford, 2010, p. 169).

Intrusion Detection System – "Hardware or software that gathers and analyzes information from various areas within a computer or a network to identify possible security breaches, which include both intrusions and misuse" (NIST, 2013, p.104).

Intrusion Prevention System – "Systems that can detect and attempt to stop an intrusive activity, ideally before it reaches its target" (NIST, 2013, p. 105).

Malicious Insider – "Is an insider who has malicious intent that acts against the best interests of the organization" (Santos et al., 2012, p. 331).

Security Event – "Any observable security occurrence in a system network" (NIST, 2012, p. 6).

Security Incident – "A violation or imminent threat of violation of a computer security policy, acceptable use policy, or standard security practice" (NIST, 2012, p. 6).

Security Information and Event Management (SIEM) Tool – "Application that provides the ability to gather security data from information system components and present that data as actionable information via a single interface" (NIST, 2013, p. 177).

Threat – "Any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service (NIST, 2013, p. 198).

Summary

The purpose of chapter one was to introduce this study by identifying the research problem, questions, barriers, issues, and limitations. In the next chapter, a comprehensive review of literature is presented. The research problem of the study was the prevalent challenge faced within the cybersecurity industry when detecting potentially malicious insider cyber threats, to enable visualization of those threats as they occur. Literature reviews to support the stated research problem and need for this study are detailed within the next chapter.

Chapter one also presented the research goals and research questions. The main goal of this study was to develop and validate, using SMEs, a cyber insider threat dashboard visualization prototype. The prototype was used in an experimental study that aimed to assess the perceived effectiveness of enhancing the presentation of complex data correlations when mitigating malicious insiders cyber threats. Literature to support this goal was presented (Albanese, Pugliese, & Subrahmanian, 2013); Boukri & Chaoui, 2015); Dork et al., 2011); Greitzer & Hohimer, 2011); Legg et al., 2015); Shneiderman et al., 2010); as well as Shneiderman & Plaisant, 2015). The five specific goals of this study were also discussed. Based on prior literature intrusion detection and prevention systems

may be insufficiently designed so they may not be capable of identifying malicious insiders cyber threats (Agrafiotis et al., 2015; Spathoulas & Katsikas, 2010). Identifying anomalous activities amidst appropriate activities pose the potential difficulty of being able to identify legitimate anomalies within the data presented by Shneiderman and Plaisant (2015). This study added to the body of knowledge by applying a novel approach to insider threat identification, new insight into enhancing mitigations by utilizing appropriate visualization techniques may be applied in order to detect malicious cyber insiders.

Lastly, chapter one continued by discussing the barriers, issues and potential mitigations for each. The limitations and delimitations of this study were also discussed within the presented barriers and issues section. The chapter concluded with a list of definitions of terms that was used throughout this study and any applicable acronyms.

Chapter 2

Review of Literature

Introduction

A literature review is presented in this chapter to provide an analysis of the relevant literature pertaining to insider threat detection and visualization. According to Webster and Watson (2002), an effective literature review is essential to creating a firm foundation for advancing knowledge, since it facilitates theory development and uncovers areas needing additional work. A literature review should also be objective and gather information on a particular subject from many sources, in order to support the newly contributed insight (Ramdhani, Ramdhani, & Amin, 2014). Thus, the presented literature review within this chapter displayed objective information gathering in relation to the need for this work. This examination consists of an extensive search performed within IS, InfoSec, HCI and medical literature. Quality literature reviews have structure, form, and is structured around major themes or concepts that emerge as the author examines and reviews the literature (Levy & Ellis, 2006). As a result of the literature review performed, constructs relevant to the visualization of insider threats for identifying anomalous activities of malicious cyber insiders were identified as: cybersecurity, cyber analysts, intrusion detection, insider threat analytic variables, information visualization and IS effectiveness. An extensive review of these constructs was preformed to determine established knowledge within these identified areas needing

additional investigation. The results of this literature review relevant to these constructs are later presented.

Cybersecurity

Cybersecurity is defined as "prevention of damage to, unauthorized use of, exploitation of, and, if needed, the restoration of electronic information and communications systems to ensure confidentiality, integrity, and availability" (Axelrod, 2006). It is commonly used to refer to the protection of devices connected to the Internet (Addae, Radenkovic, Sun, & Towey, 2016). The term 'cybersecurity' is often used interchangeably with the term information systems security, though these terms greatly differ (Solms & Niekerk, 2013). Cybersecurity is not just the protection of cyberspace; it is also the protection of those who function in cyberspace and any of their assets that can be reached via cyberspace (Thomson & Solms, 2005).

Cybersecurity is about the protection of the assets and people using resources in cyberspace and any other assets these assets may be tangible or intangible and including those belonging to society in general (Solms & Niekerk, 2013). Cybersecurity issues are different from any issues faced before (Harknett & Stever, 2011). Cybersecurity issues involve human intelligence and the exploits or vulnerabilities are created to defy and change the rules of the systems they target (Toecker, 2014). Problems within cybersecurity do not fit the traditional security variables, since problems within cybersecurity tend to also be a strategic issue and not just a compliance issue (Bissell, 2013). There is also a profuse amount of non-geographical data and an additional layer of

territorial division of responsibilities that adds to the difficulties of mitigating cyber insider threats (Harknett & Stever, 2011).

Cybersecurity problems became prominent following the attacks on September 11, 2001, after which President Bush created the U.S. Department of Homeland Security (Harknett & Stever, 2011). The Bush Administration unveiled the National Strategy for Securing Cyberspace (NSSC) in 2003, a prominent effort to address the nation's cybersecurity problems (Harknett & Stever, 2011). Cybersecurity issues persisted into 2009 and President Obama declared cybersecurity defense as a significant national security interest that the U.S. government was not adequately prepared to counter (Sherman, 2013). To understand cybersecurity it is important to also understand cyberspace.

Cyberspace consists of a growing number of connected computers, with a global community of individual users, and a constantly evolving set of technologies. Cybersecurity professionals actively work towards ensuring confidentiality, integrity, and availability for computers within this space (Miller & Murphy, 2009). With increased reliance on cyberspace, vulnerabilities have also grown (Warkentin & Willison, 2009). Anderson and Agarwal (2010) noted that cybersecurity issues are contingent upon securing cyberspace, and this issue is comparable to environmental and health issues. A defining characteristic of cybersecurity is that all assets may contain data that needs to be protected, since the vulnerabilities that exist are a result of information and communication technologies, all of which make up cyberspace (Solms & Niekerk, 2013). Caputo et al. (2009) concluded they were unable to determine a specific technique for

determining insider misuse, but by providing a common operational overview of cyberspace they have provided valuable insights for mitigating cybersecurity problems.

Table 1

Summary of Cybersecurity Literature

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Addae, Radenkovic, Sun, & Towey, 2016	Empirical study via survey	174 online participants	Survey using a five-point Likert scale	Augmented behavioral research model that introduces attitude to personal data as a determinant of cybersecurity behavior.
Anderson & Agarwal, 2010	Empirical study via survey	594 home computer users	Survey and experiment using a seven-point Likert scale	Behavioral intentions to secure one's own computer and to secure the Internet is formed by a combination of cognitive, social, and psychological components.
Axelrod, 2006	Literature review and analysis		IT process and control requirements	Enterprises must take on responsibility for protection within their perimeters.

Table 1

Summary of Cybersecurity Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Bissell, 2013	Literature review and synthesis		Cyberattacks, cyberinsurace	The development of a cybersecurity roadmap enabled organizations to develop a much clearer picture of its current status and gain a better understanding of its strengths and gaps.
Caputo et al., 2009	Empirical study via experiment	50 employees at MITRE (Management, technical, and administrative staff)	A study laptop running software that monitored their information-use behavior	There isn't one behavior that distinguishes malicious users from benign ones. The most valuable way to tackle insider threats is to cast a wide net and strategically evaluate behaviors to identify misuse.
Harknett & Stever, 2011	Policy review and analysis	U.S. government cybersecurity policy releases from 2002 to 2011	Cybersecurity policy	Based on the national strategy for securing cyberspace the Goal of securing cyberspace was occasionally achieved in moderate technical, tactical, and operational advances.

Table 1

Summary of Cybersecurity Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Miller & Murphy, 2009	Theoretical		Cybersecurity review and problems	The cybersecurity problem required new models of cooperation and collaboration among nations.
Sherman, 2013	Literature review and analysis	58 U.S. government releases on cybersecurity	Cybersecurity defense, cybersecurity strategy	Senior policymakers must commit the resources to build and nurture a highly skilled cyber workforce capable of overcoming cyber threats and vulnerabilities.
Solms & Niekerk, 2013	Literature review and Synthesis		Cybersecurity and information security	Cybersecurity, differs from information security; cybersecurity is not only the protection of cyberspace itself, but also the protection of those that function in cyberspace and any of their assets.

Table 1

Summary of Cybersecurity Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Thomson & Solms, 2005	Literature review and synthesis		Information security obedience, corporate governance, information security, corporate culture	Information Security Obedience is the solution to ensuring proper information security behavior.
Toecker, 2014	Theoretical		security control systems	Recommendations to provide efficient risk reduction from cybersecurity events, answering the question "Where should I put my next dollar in order to get the biggest cybersecurity improvement?"
Warkentin & Willison, 2009	Literature review and synthesis		Insider threat	There is a need to understand and address the various risks to the security of the IS on which we depend.

Cyber Analysts

Organizations rely on skilled analysts to make critical decisions pertaining to: threats, vulnerabilities, and network performance (Staheli, Yu, Crouser, Damodaran, Nam, O'Gwynn, McKenna, & Harrison, 2014). Cyber analysts examine people within

organizations, internal, external incidents, within varying applications, locations, and dates to uncover potentially imminent threats (Gorg et al., 2013). They have the unsavory job of analyzing a profuse amount of alerts whose performance characteristics are often either unknown or uncertain (Walton, Maguire, & Chen, 2015). Analysts struggle with processing large volumes of data and providing valuable insights (Shneiderman & Plaisant, 2015). In many cases the overwhelming amount of data leaves cyber analysts unable to formulate effective remediation plans (Arias-Hernández et al., 2011). Caputo et al. (2009) found that two cyber analysts could effectively review 23 users per day using a developed insider threat detection system. However, they still struggled to develop a technique which aided in efficient identifying, with significant reduction in false positives, of malicious insiders.

Visualizations benefit cyber analysts attempting to identify complex problems. Visualizations help cyber analysts to both identify problems, potentially malicious insiders, and to work visually towards finding solutions (Fink, North, Endert, & Rose, 2009). Goodall (2007) evaluated user performance when using a traditional application versus one that utilizes visualizations for network packet analysis and found that users discovered more insights when utilizing a visualization tool. Visualizing complex issues using simple and intuitive methods so patterns can be quickly recognized assisted in overcoming cybersecurity problems (Choi, Lee, & Kim, 2009). Though Goodall (2007) found visualizations to be beneficial, his study focused only on network analysts and suggested the methods be applied using cybersecurity applications.

Table 2

Summary of Cyber Analysts Literature

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Arias-Hernández et al., 2011	Theoretical		Cognitive science, visual analytics, HCI	For analysts to overcome visual analytics challenges it would be best achieved with the active involvement of HCI researchers and practitioners.
Caputo et al., 2009	Empirical study via experiment	50 employees holding various positions at MITRE (Management, technical, and administrative staff)	Study laptop running software that monitored information-user behavior	There isn't one behavior that distinguishes malicious users from benign ones. The most valuable way to tackle insider threats is to cast a wide net and strategically evaluate behaviors to identify misuse.
Choi, Lee, & Kim, 2009	Empirical study via experiment	A real-life Internet attack traffic trace (Network packets)	Traffic flow generator	The development of parallel coordinate attack visualization (PCAV) is a real-time visualization system for detecting anomalies from Internet attacks.
Fink, North, Endert, & Rose, 2009	Empirical study via experiment	8 cyber analysts at a laboratory	Visualization	Designed usable workspaces for cyber analysts.

Table 2

Summary of Cyber Analysts Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Goodall, 2007	Literature review and synthesis		Visualization	Examination of VizSec literature attempting to solve the problems of computer security through enabling humans through information visualization.
Gorg et al., 2013	Empirical study via experiment	5 Student teams at Mercyhurst College for 10 weeks	Visual analytics	Visual analytics succeeds only if developers fully understand the unique demands of analysis and the way that analysts approach their work.
Shneiderman & Plaisant, 2015	Literature review and synthesis		Analytic-focusing	Identification of 10 analytic focusing strategies to sharpen analytic processes and enable users to deal with larger datasets.
Staheli, Yu, Crouser, Damodaran, Nam, O'Gwynn, McKenna, & Harrison, 2014	Empirical study via secondary data and survey	Surveyed 130 papers from the past 10 years of VizSec proceedings	Visualization	Identified existing methodological gaps in evaluating visualization in cyber security, and suggested potential avenues for future research.

Table 2

Summary of Cyber Analysts Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Walton, Maguire, & Chen, 2015	Empirical study via case study	Selection of visualizations from the system, CITD (Corporate Insider Threat Detection) Dashboard	Dashboard Visualization	Introduced a visual analytics loop for protective monitoring in cybersecurity applications, and a prototype tool demonstrating an example implementation of the approach.

Intrusion and Anomaly Detection

Intrusion detection systems are the equivalent of burglar alarms within cybersecurity and anomaly detection systems are a subset of IDSs that specialize in discovering unknown attacks (Patcha & Park, 2007). Although anomaly detection stemmed from IDSs, the goal of an anomaly detection system is to detect new or unknown attacks (Yu, 2012). Anomaly detection relies on identifying normal activities and reporting against deviations from identified normal activities (Cao, Li, Coleman, Belatreche, & McGinnity, 2015). Anomaly detection is based on event correlation techniques that can be categorized and utilized to infer threats through correlation analysis (Ten, 2010). Common types of IDSs are signature based or anomaly based (Ye, Emran, Chen, & Vilbert, 2002). Commonly used IDSs are often signature based and require constant updates of rules and known attacks to stay effective (Patcha & Park, 2007). Signature based detection is reactive and with today's advanced threats they are

seemingly outdated (Jackson, 2012). Signature based anomaly detection uses intrusion signatures that have to be manually added as profiles of intrusion characteristics, if an intrusion signature is present then and intrusion alert is triggered (Ye et al., 2002).

Anomalies are patterns of interest that do not conform to normal behaviors detection techniques use identified data to develop a baseline of normal activities (Chouhan & Richhariya, 2015). Some anomaly detection approaches use the identification of a score, to indicate "the degree of irregularity of a specific event", when activities result in a score that exceeds the established baseline of normal activity, then the occurrence was flagged as an anomaly (Garcia-Teodoro, Diaz-Verdejo, Maciá-Fernández, & Vázquez, 2009, p. 20). Unknown attacks are detected by creating a baseline of 'normal' activities and if any activities deviate from the baseline then it is identified as anomalous and potentially malicious activity (Patcha & Park, 2007). Thus, activities that exceed the baseline as well as activities that are significantly below the baseline can initiate further investigations by cyber analysts. The baseline running an average for each pertinent variable, if the data sets activities are above average it was deemed above the baseline and vice versa. The baseline may be organizational-dependent and could rely upon factors like the number of employees, type of data being collected, number of data sources, etc. (Legg et al., 2015). Like organizations, individuals may be drastically different, however, there are established 'normal' baselines for each vital sign in relation to an individual infant or adult. For instance, the vitals of a healthy adult are depicted using an Electrocardiogram (EKG) view figure 3, presenting a standard set of vital signs: "blood pressure: 90/60 mm/Hg to 120/80 mm/Hg, respiratory rate: 12–20 breaths per minute, pulse rate: 60–100 beats per minute, and temperature: 36°C–37.4°C" (Mok et al.,

2015, p. 98). Hence, comparable cybersecurity vital signs may be established.

Figure 3: Standard EKG Monitor

Table 3

Summary of Intrusion and Anomaly Detection Literature

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Cao, Li, Coleman, Belatreche, & McGinnity, 2015	Empirical study via experiment	Real market data of seven representative stocks: Google, Microsoft, Intel, Apple ARM, BARCLAYS, and Vodafone	Detection algorithm, adaptive hidden Markov model with anomaly states (HMMAS)	AHMMAS performs better in terms of the area under ROC curve and the F-measure, respectively.

Table 3

Summary of Intrusion and Anomaly Detection Literature (continued)

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Chouhan & Richhariya, 2015	Literature review and analysis		Anomaly detection algorithm	Summarized anomaly detection techniques along with various research directions.
Garcia-Teodoro, Diaz-Verdejo, Maciá-Fernández, & Vázquez, 2009	Literature review and analysis		Anomaly-based network intrusion detection systems (A-NIDS)	Faster and more effective countermeasures are needed to cope with the ever-growing number of detected attacks.
Legg et al., 2015	Empirical Investigation via experiment	Ten scenarios within a prototype system	Anomaly detection	An approach for insider threat detection based on organizational log data, the system generates user and role-based profiles that can describe the full extent of activities that users perform within the organization.
Mok et al., 2015	Literature review and synthesis	A publication search between 1990 to November 2012	Vital Signs	Observation chart designs together with proper training can most likely improve the detection of deteriorating vital signs.

Table 3

Summary of Intrusion and Anomaly Detection Literature (continued)

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Chouhan & Richhariya, 2015	Literature review and analysis		Anomaly detection algorithm	Summarized anomaly detection techniques along with various research directions.
Garcia-Teodoro, Diaz-Verdejo, Maciá-Fernández, & Vázquez, 2009	Literature review and analysis		Anomaly-based network intrusion detection systems (A-NIDS)	Faster and more effective countermeasures are needed to cope with the ever-growing number of detected attacks.
Legg et al., 2015	Empirical Investigation via experiment	Ten scenarios within a prototype system	Anomaly detection	An approach for insider threat detection based on organizational log data, the system generates user and role-based profiles that can describe the full extent of activities that users perform within the organization.
Mok et al., 2015	Literature review and synthesis	A publication search between 1990 to November 2012	Vital Signs	Observation chart designs together with proper training can most likely improve the detection of deteriorating vital signs.

Table 3

Summary of Intrusion and Anomaly Detection Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Ye, Emran, Chen, & Vilbert, 2002	Literature review and synthesis		Intrusion detection	Since intrusions may manifest more through mean shifts than through counter relationships, we can suppress noises and variations in normal activities causing counter relationships to improve the accuracy of intrusion detection.
Yu, 2012	Literature review and synthesis		Intrusion detection	The evolution of intrusion detection systems over the past two decades.

Anomaly Detection Techniques

The most important step for anomaly detection is based on the delineated data sources profiling the system and user activities. A series of mathematical inputs and outputs are utilized as predictive methods when detecting anomalies (Patan, 2015). The method for anomaly detection may vary based on the data source input. Data sources can include shell commands, system events, audit events, user keystrokes, and packages that traverse the network (Jyothsna, Prasad, & Prasad, 2011). Data is collected continuously and may be from several heterogeneous data sources (Chandola, Banerjee, & Kumar, 2009). There are various anomaly detection techniques relevant to the data source

identified, prior research denotes anomaly detection techniques as: statistical, cognition, and machine learning based, as depicted in Figure 4 (Jyothsna et al., 2011; Garcia-Teodoro et al., 2009).

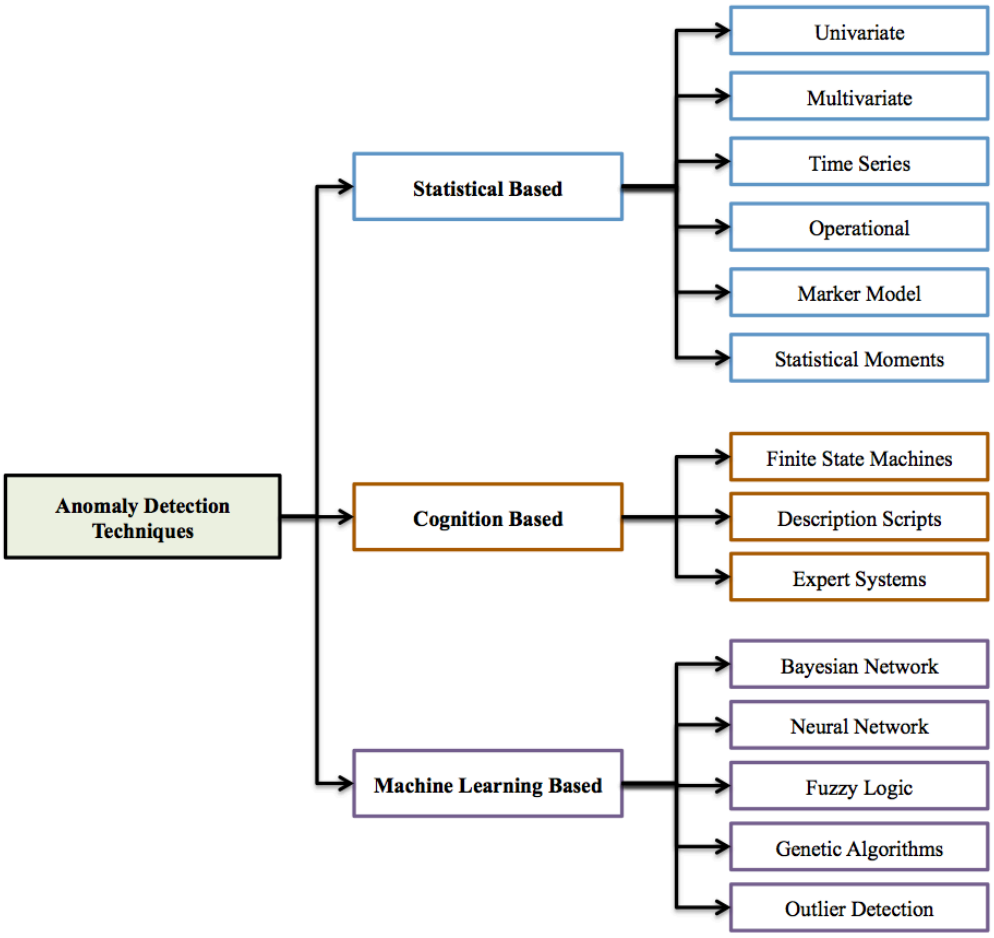


Figure 4: Anomaly Detection Techniques
(Garcia-Teodoro et al., 2009; Jyothsna et al., 2011)

Statistical based anomaly detection techniques capture the network traffic and develop a behavioral profile. Using statistical properties such as mean and variance to build a profile of normal activities, then statistical tests are used to determine significant deviations from the normal profile (Qayyum, Islam, & Jamil, 2005). Within statistical anomaly detection there are two steps involved: first “normal behavior” is characterized, then a time frame where behavior does not seem to be normal is determined (Wang &

Paschalidis, 2015). There are several models within the statistical based techniques this includes: univariate, multivariate, time series, operational, Markov or Marker models, and statistical moments (Jyothsna et al., 2011). Univariate models utilize a single metric as well as multivariate models that use correlations of two or more metrics to determine deviations (Jyothsna et al., 2011). Time series models uses an interval timer with event counters or resource measures, then consider the order and time frames of each activity and their value, so if at a given time the traffic observed is too low it may be identified as an anomaly (Qayyum et al., 2005). The operational model also referred to as the threshold metric is based on cardinality of events that happen over a period of time, by counting events as they occur, then triggering an alert if the number of events is higher or lower than the specified thresholds (Jyothsna et al., 2011). The marker model is also known as the 'Markov model' which was broken into two approaches: Markov chains and hidden Markov models (Garcia-Teodoro et al., 2009).

Cao et al., (2015) identified that hidden Markov models (HMM) using one time series data was insufficient and developed an adaptive HMM (AHMAS) with anomaly states to detect anomalies based on a sequence of data and not a single value at a point in time. Within the Markov model, a Markov chain is a set of states connected by transition probabilities anomalies are detected by comparing the associated probability with the observed sequence (Garcia-Teodoro et al., 2009). Statistical moments term all identified correlations as 'moments', if an event occurs above or below an identified moment the activity is deemed to be anomalous (Jyothsna et al., 2011). The moment is a mean or standard deviation of the correlations, the system determines the confidence interval

based on observed user data and does not require prior knowledge of normal activities (Qayyum et al., 2005).

Cognition based anomaly detection consists of finite state machine models, description scripts, and adept or expert systems (Jyothsna et al., 2011). Cognition based anomaly detection techniques use human input to determine legitimate behaviors by having SMEs manually construct the desired model (Garcia-Teodoro et al., 2009). A finite state machine (FSM), captures actions (entry, exit, & transition action) in states, an action is a description of an activity to be performed at a given moment and contains information about the past (Jyothsna et al., 2011). Description scripts are scripting languages developed by the Intrusion Detection community to identify attacks based on sequences of specific events by describing signatures of attacks (Jyothsna et al., 2011). Expert systems classify audit data according to rules, by first identifying different attributes, then the classification parameters and procedures are determined, then finally the data is classified accordingly (Garcia-Teodoro et al., 2009).

Machine learning based anomaly detection consists of Baysian networks, generic algorithms, neural networks, fuzzy logic, and outlier detection models (Jyothsna et al., 2011). Machine learning based anomaly detection techniques are based on explicit of implicit models that allow patterns to be analyzed for categorization (Garcia-Teodoro et al., 2009). Bayesian networks focus on identifying problematic relationships based on integrated prior knowledge (Garcia-Teodoro et al., 2009). The Baysian network model is a mathematical framework for combining information to perform estimation by combining known information to postulate unknown information (Zaknich, 1998). By attaining knowledge through learning neural network models mimic the brain (Garcia-

Teodoro et al., 2009). Instead of utilizing precise rules, by using approximate rules fuzzy logic can improve detection accuracy (Xu, You, & Liu, 2005). Fuzzy logic can be used to match any input or output of data by focusing on a range of variability and not just an exact outcome this aids with understanding vague or ambiguous information (Dutta et al., 2013). Genetic algorithms identify deviations with no prior knowledge of the system behaviors by using techniques stemming from biology, including inheritance, mutation and selection (Garcia-Teodoro et al., 2009). For instance, based on genetic algorithms the use of a negative selection algorithm that filters out bad solutions tends to reduce the rate of false positives within anomaly detection (Jyothsna et al., 2011). Outliers are unusual activities that defer from normal activities (Zhang & Zulkernine, 2006). Outlier detection consists of grouping data observations according to a given similarity, based on identified similarities each data point is grouped. Points that fall outside of the grouped clusters are classified as outliers (Garcia-Teodoro et al., 2009).

Patcha and Park (2007) found that "today's intrusion detection approaches will not be able to adequately protect tomorrow's networks against intrusions and attacks" (p. 3465). Therefore, anomaly detection methods need to be advanced to address this problem. Jyothsna et al. (2011) denoted that identifying features to characterize user and system patterns would be the best way to clearly distinguish anomalous activities. Wang and Paschalidis (2015) found that existing anomaly detection methods tend to focus on stationary assumptions, suggesting there is no change over time, and therefore proposed a more robust model for network traffic analysis. Hence, as the anomaly detection methods evolve, the systems applying these methods need to be advanced as well. This study developed and validated cyber vital signs to be presented within a cybersecurity insider

threat dashboard visualization prototype, by incorporating a novel approach to anomaly detection using SME identified visualization techniques.

Table 4

Summary of Anomaly Detection Techniques Literature

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Cao et al., 2015	Empirical study via experiment	Real market data of seven representative stocks: Google, Microsoft, Intel, and Apple from NASDAQ, ARM, BARCLAYS, and Vodafone against OCSVM, kNN, and GMM models	Detection algorithm, adaptive hidden Markov model with anomaly states (HMMAS)	The comparison of proposed approach with other benchmark models, OCSVM, kNN, and GMM, has shown that the AHMMAS performs better in terms of the area under ROC curve and the F-measure, respectively.
Chandola, Banerjee, & Kumar, 2009	Literature review and synthesis		Anomaly detection	Current research has been done in an unstructured fashion, without relying on a unified notion of anomalies, which makes the job of providing a theoretical understanding of the anomaly detection problem.

Table 4

Summary of Anomaly Detection Techniques Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Dutta et al., 2013	Empirical study via experiment	Using one investigator data was captured from the experiment: (top to bottom): angle, footplate force, ECG, blood pressure.	Fuzzy logic, decision support	Computational approach to monitoring vital signs since the purpose of such alerts is to provide decision support inputs to careers, to prompt closer observations or direct interventions to be performed to help the subjects of care.
Garcia-Teodoro et al., 2009	Literature review and analysis		Anomaly-based network intrusion detection systems (A-NIDS)	Faster and more effective countermeasures are needed to cope with the ever-growing number of detected attacks.
Jyothsna, Prasad, & Prasad, 2011	Literature review and analysis		Intrusion, anomaly-based, and signature based detection	Faster and more effective countermeasures are needed to cope up with the attacks ever-growing.
Patan, 2015	Empirical study via experiment	Signals from a boiler	A recurrent neural network with open-loop control system	In spite of proper work of the developed control algorithms, selecting control parameters were still a challenge.

Table 4

Summary of Anomaly Detection Techniques Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Patcha & Park, 2007	Literature review and analysis		Anomaly detection	Traditional intrusion detection systems have not adapted adequately to new networking paradigms like wireless and mobile networks nor have they scaled to meet the requirements posed by high-speed networks.
Qayyum, Islam, & Jamil, 2005	Literature review and analysis		Anomaly detection, intrusion detection, security threats	Advantages and disadvantages of various techniques for anomaly detection.
Wang & Paschalidis, 2015	Empirical study via experiment		Network anomaly detection	The statistical properties of normal traffic are time-varying for most actual communication networks.
Xu, You, & Liu, 2005	Empirical study via experiment	Mackey-Glass Time series and a publicly available data set	Negative selection, genetic algorithm, fuzzy logic	Combined negative selection and genetic algorithm to develop a novel fuzzy rules based approach to system performance detection.

Table 4

Summary of Anomaly Detection Techniques Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Zaknich, 1998	Literature review and analysis		modified probabilistic neural network, general regression neural network	Introduced the MPNN and described its relation to Specht's GRNN, MPNN has some advantages over other neural networks in nonlinear signal processing applications.
Zhang & Zulkernine, 2006	Empirical study via experiment	Attack data	Network Intrusion Detection Systems	A new framework for unsupervised anomaly NIDS based on the outlier detection technique in random forests algorithm.

Insider Threat Analytic Indicators

An insider threat analytic indicator, or for the purpose of this study, cyber visualization variables, are defined as outputs that may indicate an insider threat and prompts for further analysis (Casey, 2015). Cappelli, Moore, and Trzeciak (2012) suggested that automated and manual countermeasures may be utilized for mitigation based on indicators that could suggest an increased risk. Alahmadi, Legg, and Nurse (2015) proposed that a more comprehensive analytic approach incorporating a diverse set of data with sources from technological, behavioral, and physiological monitoring is more effective in recognition, detection and response to insider threats. Casey, Koeberl,

and Vishik (2010) stressed the importance of threat agent analysis to help form a "coherent picture of the threat space and priorities of remediation" (p. 1).

Casey (2007) developed a model for simplified depiction of threat agents. The Threat Agent Library (TAL) provides a consistent reference library describing the human agents involved in IT systems and that could pose threats to such kind of systems, although not limited to insider threats (Nostro, Ceccarelli, Bondavalli, & Brancati, 2013). Carcary (2013) noted the TAL as applicable for effective risk management. Ceccarelli et al. (2015) notes that identifying possible human agents that could pose a threat to IS can be very challenging, and the TAL provides a standardized set of agents. The TAL allowed for easy classification of insider threat agents and variables (see Figure 4).

Insider attacks are difficult to detect, and many attackers operate within their granted restrictions, thus, identification of potential threat agents aided in insider threat mitigation (Nostro et al., 2013). In this study, cyber visualization variables were identified using SMEs. In order to identify insider threats and the variables associated with potential threats, an understanding of who a potential threat agent is needed (Callegati, Giallorenzo, Melis, & Prandini, 2016). This study presents a novel way of addressing the prevailing problem faced when detecting malicious cyber insiders. Consequently, Figure 5 presents insider threat agent types as malicious, non-malicious, or either. This study focused on potential threats that are highlighted in Figure 5, these threat agents are identified as 'malicious or either'. The threat agents within the scope of this study are identified as: vendors, partners, irrational individuals, thief, disgruntled individuals, activists, terrorists, organized crime, competitors, and nation states. Depending upon the type of threat agent correlated attack types are depicted. Nostro et al.

(2013) utilized this predefined library of insiders in order to efficiently perform tasks relevant to the focus of their study and show completeness of the assessment of insiders. The same approach has been utilized for the purpose of this study.

Threat Agent Library (TAL)										
			Attack Types							
			Accidental Leak	Espionage	Financial Fraud	Misuse	Opportunistic Data Theft	Physical Theft	Product Alteration	Sabotage	Violence
Insider Threat Agent Types	Non-Malicious	Reckless Individual	x			x			x		
		Untrained / Distracted Individual	x			x			x		
		Outward Sympathizer	x			x					
	Either	Vendor	x	x	x	x	x		x		
		Partner	x	x	x	x	x		x		
	Malicious	Irrational Individual	x			x		x		x	x
		Thief		x	x		x	x			
		Disgruntled Individual	x	x	x	x	x	x	x	x	x
		Activist		x		x	x		x	x	
		Terrorist						x		x	x
		Organized Crime		x	x		x	x	x		
		Competitor		x			x		x	x	
		Nation State		x			x		x	x	

Figure 5: Threat Agent Library (TAL)(Casey, 2007)

Casey (2007) also presented analytics based on attack types, providing a comprehensive list of analytic indicators. For the purpose of this study, these analytic indicators provided a foundation as the initial list of critical cyber visualization variables. Based on the variable and type of analytic indicator the determination of whether the variable is applicable when identifying potentially malicious cyber insiders can be determined. Data associated with the identified variable needs to be depicted in the developed cyber visualization. Wang and Jones (2017) suggested that flows, logs and system events (alerts) are typically used for intrusion detection. Cappelli et al. (2012) identified that log traffic may be inspected for indicators of "suspicious access, large file transfers, suspicious email traffic, after-hours access, or use of removable media by

resigning employees" (p. 93). Brdiczka et al. (2012) denoted that model variables can be associated with observable indicators. Figure 6 depicts an initial set of variables for the identification of potentially malicious insider cyber threats, and the associated data.

Category	Analytic Variables	Data Types				
Activity-Based Analytics		**Alerts**	**Content**	**Flows**	**Logs**	**Identity**
System	Authentication and Authorization Failure	x			x	x
	Changes in Data Access Patterns				x	x
	Access Inconsistent With User Class				x	x
	Changes in Network Patterns			x		x
	Network Patterns Inconsistent with User Class			x		x
	Data Exfiltration	x	x	x	x	x
	Unauthorized Data Access Methods	x			x	x
	Privilege Change	x		x	x	x
	Erroneous Defensive Posture Changes	x		x	x	x
	Improper Command Usage	x			x	x
	Knowledge Access		x	x	x	x
	Audit Log Modification	x			x	x
Facility	Time of Access Pattern Changes	x		x	x	x
	Locality of Access Pattern Changes	x		x	x	x
Business Capabilities	Failure Correlation		x		x	x
	Malware Deployment	x			x	x
	Deletion or Modification of Data or Infrastructure	x			x	x
	Analysis of Competitor		x			x
	Analysis of Public Media		x			x
	Attribution of Disclosure		x		x	x
	Retrieval			x	x	x
Content-Based Analytics		**Alerts**	**Content**	**Flows**	**Logs**	**Identity**
Social	Disregard		x			x
	Personal Inflexibility		x			x
	Unusual Contacts		x			x
	Unusual Business Travel		x			x
	Unusual Personal Travel		x			x
	Unauthorized or Inappropriate Associations		x	x		x
	Withdrawal		x	x		x
	Workplace Events		x			x
	Workplace Satisfaction		x			x
Health	Mental instability		x			x
	Impulse Control		x			x
Human Resources	Major Life Event		x			x
	Complaints Against the User		x			x
	Negative Reviews		x			x
Inferential Analytics		**Alerts**	**Content**	**Flows**	**Logs**	**Identity**
Financial	Observed Temporal Change in Means		x			x
	Observed Change in Means Relative to Peers		x			x
	Financial Reporting		x			x
Security	Change in Violation Patterns	x			x	x
	Duration and Regularity of Security Events	x			x	x
	Unauthorized or Inappropriate Use of Tools	x			x	x
Criminal	Restraining Orders		x		x	x
	Wage Garnishments, etc.		x		x	x
	Violence Outside Workplace		x		x	x
	Recent Increase in Criminal Events	x	x		x	x

Figure 6: Initial Analytic Variables and Data Types (Casey, 2015)

Table 5

Summary of Insider Threat Analytic Indicators Literature

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Alahmadi, Legg, & Nurse, 2015	Empirical study via experiment	1000 randomly selected websites	Website–OCEAN personality correlation tool, insider threat application	An individuals browsing interests can lead to inferences about their personality traits, and if monitored can be utilized to identify potential insider threat.
Brdiczka, Price, Shen, Patil, Chow, Bart, & Ducheneaut, 2012	Empirical study via experiment	350,000 multi-player online game characters observed over a period of 6 months	Structural anomaly detection, psychological context modeling	An approach for proactive detection of insider threats by combining structural anomaly detection, information networks, and psychological profiling of individuals.
Callegati, Giallorenzo, Melis, & Prandini, 2016	Literature review and analysis		Threat agents	Developing a Federated marketplace of services called SMAll, aimed at harmonizing data flows and service invocations.
Cappelli, Moore, & Trzeciak, 2012	Literature review and synthesis	CERT Database	Theft, sabotage, fraud	Identification of types of insiders with varying motives of why they perform malicious activities.

Table 5

Summary of Insider Threat Analytic Indicators Literature (continued)

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Carcary, 2013	Literature review and analysis		Risk management	A new IT Risk Management maturity model.
Casey, Koeberl, & Vishik, 2010	Theoretical		Threat agents, threat agent library (TAL)	Standardization of threat agents.
Ceccarelli, Montecchi, Brancati, Lollini, Marguglio, & Bondavalli, 2015	Empirical study via experiment	Data for four different biometric traits applied within a math lab	Context aware security by hierarchical multilevel architectures (CASHMA), TAL, ADVISE (ADversary VIew Security Evaluation) modeling formalism	Utilized biometrics to define a protocol for continuous authentication that improves security and usability of user sessions.
Nostro, Ceccarelli, Bondavalli, & Brancati, 2013	Empirical study via case study	A service oriented system	TAL, ADVISE (ADversary VIew Security Evaluation) modeling formalism	Defined the motivations and targets of insiders, investigate the likeliness and severity of potential violations, and identified appropriate countermeasures.

Table 5

Summary of Insider Threat Analytic Indicators Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Wang & Jones, 2017	Literature review and synthesis		Big data analytics, anomaly detection, misuse detection	Methods for network intrusion detection, stream data characteristics and stream processing systems, feature extraction and data reduction, conventional data mining and machine learning, deep learning, and Big Data analytics in network intrusion detection.

Information Visualization

Information visualization is defined as communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014). Using effective visualization interfaces within cybersecurity allows the ability to understand anomalies using data from complex cyber networks (Inibhunu et al., 2016). Rohrer and Swing (1997) attested that when dealing with unstructured data utilizing the proper visual mapping would be essential for producing effective visualizations. However, Fink et al. (2009) noted that most cyber analysis still utilized archaic command-line tools that are ineffective at presenting large volumes of rapidly moving data. Choi, Lee, and Kim (2009) proposed a new visualization technique to detect anomalies from Internet attacks since traditional IDSs were not able to detect unknown attacks without triggering a significant amount of false positive alerts. Visualizations should be useful in presenting

unstructured data, and not only be valuable in presenting an outcome once it is already known (Kandel, Paepcke, Hellerstein, & Heer, 2012).

Inibhunu et al. (2016) argued that using level of detail viewing would increase performance and decrease information overload, facilitating effective mitigation decisions within cybersecurity. Fink et al. (2009) believed large, high-resolution displays with interoperable, flexible, and compelling visualization tools are core components of a usable workspace for cyber analysts. However, they did not examine how the actual information displayed using the visualizations should be enhanced to support cyber analysts. Since the focus of the visualizations that were being developed were not being tailored towards the needs of a cybersecurity analyst. After surveying 32 analysts, Kandel et al. (2012) found that analysts would like visualizations that allow them to apply advanced analytics routines, explore models, and visualize the output. McKenna et al. (2015) identified that more cyber visualizations need to be developed from the perspective of cybersecurity analysts.

McKenna, Mazur, Agutter, and Meyer (2014) provided a framework for actionable guidance throughout the visualization design process by focusing on two models for visualization design, decision models to capture the rational behind the decisions designers make, and process models capturing the actions made. This allowed visualization designers to verify and validate design decisions, allowing the final design to be more applicable to the real world. McKenna et al. (2015) assessed that only 40% of visualization research gathered user needs prior to developing a tool, thus, they assessed user needs within cybersecurity, developed, and then validated the visualization. However, the validation was only performed with one cybersecurity analyst. Walton et al.

(2015) noted that there are multiple shortcomings of visual analytics workflows proposed a scalable loop for continuous improvements of the models being used to contribute to the development of more proactive and visual models. While McKenna et al. (2015) as well as Walton et al. (2015) approached the cyber visualization problem from a user focused perspective, the resultant models developed do not hone in on the user of interest or potential malicious insider. Therefore, the analysts needed to determine this based on the presented visualizations.

Table 6

Summary of Information Visualization Literature

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Choi, Lee, & Kim, 2009	Empirical study via experiment	A real-life Internet attack traffic trace (Network packets)	Traffic flow generator	The development of parallel coordinate attack visualization (PCAV) a real-time visualization system for anomaly detection.
Fink, North, Endert, & Rose, 2009	Empirical study via experiment	8 cyber analysts at a government laboratory	Visualization	Designed usable workspaces for cyber analysts.
Inibhunu et al., 2016	Empirical study via data analysis	Millions of records from firewall and IDS logs for a fictitious organization	Cyber situation awareness	Level of detail viewing can greatly increase human performance by mitigating information overload.

Table 6

Summary of Information Visualization Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Kandel, Paepcke, Hellerstein, & Heer, 2012	Empirical study via interviews	35 data analysts in a commercial organization	Visualization	As the scale and diversity of data sources increases within enterprises, there is an opportunity for visual analytic tools to improve the quality of analysis and the speed at which it takes place.
McKenna et al., 2015	Literature review and analysis		Design activity framework, cognitive task analysis	Personas, their behaviors, and their knowledge played a critical role in helping to decide which users and needs to target in the visualization design process.
Roberts et al., 2014	Theoretical		Visualization	Visualization researchers should exploit the experience gained over the last two decades in virtual reality (VR) research, while continuing to apply VR technology to visualization systems.

Table 6

Summary of Information Visualization Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Rohrer & Swing, 1997	Literature review and analysis		Information visualization	Interlinking visualization components with other Web media and data is useful since it offers fairly seamless integration of related information for end users.
Walton et al., 2015	Empirical study via case study	Selection of visualizations from the system, CITD (Corporate Insider Threat Detection) Dashboard	Dashboard visualization	Introduced a visual analytics loop for protective monitoring in cybersecurity applications, and a prototype tool demonstrating an example implementation of the approach.

IS Effectiveness

Though prior studies have struggled to define effectiveness as well as how to accurately measure effectiveness (Bailey & Pearson, 1983; Doll, Xia, & Torkzadeh, 1994; Lee, Kim, & Lee, 1995). Effectiveness has been defined as the extent to which the IS contributes to the accomplishment of objectives (Kim, 1989). Levy (2006) denoted that the complete measurement of IS effectiveness must include measures of the causal factors or 'values' as well as the end result construct or 'user satisfaction' (p. 60). Levy, Murphy, and Zanakis (2010) emphasized the importance of value as an important

construct within IS research. Levy (2006) denoted that Value Theory and User Satisfaction theory suggests that values impact attitudes which impact behaviors, and in turn impacts satisfaction (p. 6). Therefore, this study addressed perceived effectiveness as a measure of value and satisfaction.

User satisfaction with an IS is related to the utilization and success of a system, and can be measured in an effort to improve the system (Bailey & Pearson, 1983). Levy (2006) indicated that satisfaction should be measured as a 'surrogate' for measurement of IS effectiveness (p. 42). Bano, Zowghi, and Rimini (2017) found that user involvement in the development process leads to higher levels of user satisfaction. Kurucay and Inan (2017) stated that user satisfaction is a strong determinant of effectiveness. User satisfaction is an important theoretical issue, however, Doll et al. (1994) argued that it is a one-dimensional construct. Bailey and Pearson (1983) argued satisfaction is a bi-dimensional attitude, thus, the intensity of an individuals reaction relative to the information requirements must also be measured (Bailey & Pearson, 1983). Bano et al. (2017) also confirmed satisfaction to be bi-dimensional, as it entails user satisfaction with their involvement process and satisfaction with the delivered product.

Value is defined as the core belief about a level of importance that the user attributes to the system (Levy, 2006; Hackney, Dooley, Levy, & Parrish, 2015; Dooley, Levy, Hackney, & Parrish, 2017). The expectancy value-theory explains motivation as a combination of the users needs and the value of the goals in the system (Sigaard & Skov, 2015). According to Sedera, Lokuge, Grover, Sarker, and Sarker (2016) for innovation to take hold within systems their needs to be increased value. For this study the developed cyber visualization prototype was considered effective when the SMEs perceive the

developed prototype as highly important and the SMEs are highly satisfied with the visualizations (Levy, 2006). Levy (2006) used a Likert-type rating scale for assessing value items. Sedera et al. (2016) utilized a 7-point Likert rating scale to assess value of enterprise systems and digital platforms, finding that digital platforms can yield innovation, only through the moderation of the enterprise system platforms. Kurucay and Inan (2017) utilized a 5-point Likert rating scale to determine student satisfaction with online learning, to gauge the effectiveness of the online course. Thus, this study utilized a 7-point Likert-type rating scale for satisfaction and value assessment.

Table 7

Summary of IS Effectiveness Literature

Study	**Methodology**	**Sample**	**Instrument or Construct**	**Main Finding or Contribution**
Bailey & Pearson, 1983	Empirical study via survey and case study	29 questionnaires ; 32 manager interviews	Survey using a seven-point scale of satisfaction	Measurement of IS user satisfaction.
Bano, Zowghi, & Rimini, 2017	Empirical study via case study interviews	Secondary data from two case studies, 12 subjects	Three-point scale of satisfaction	User satisfaction is considered to contribute to system success.
Doll, Xia, & Torkzadeh, 1994	Empirical study via survey	409 computer users	End user computing satisfaction (EUCS)	Validation that the 12-item EUCS instrument explains and measures user satisfaction.

Table 7

Summary of IS Effectiveness Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Doskey, Mazzuchi, & Sarkani, 2015	Empirical study via experiment	27 competencies compared on successful and challenged projects	Effectiveness, Bayesian belief network, SE REI	System engineering (SE) Relative Effectiveness Index (REI) model can be used to assess SE performance.
Harrati, Bouchrika, Tari, & Ladjailia, 2016	Empirical study via experiment	50 lecturers from the Computer Science and Electrical Engineering at different universities	System Usability Scale (SUS)	System Usability Scale is not adequately a standalone measure for expressing the true acceptance and satisfaction.
Kurucay & Inan, 2017	Empirical study via experiment	77 students in an online course	24 items on a five-point Likert scale	Interaction among learners in online courses lead the higher student satisfaction.
Lee, Kim, & Lee, 1995	Empirical study via case study and survey	236 end users from 11 companies	Satisfaction, EUCS	There is a strong relationship among end-user IS acceptance, IS satisfaction and job satisfaction.
Levy, 2006	Empirical study via experiment	192 undergraduate students	IS Effectiveness, LeVIS index, EUCS	Identifying and defining the relationship between value and satisfaction in order to indicate IS effectiveness.

Table 7

Summary of IS Effectiveness Literature (continued)

Study	Methodology	Sample	Instrument or Construct	Main Finding or Contribution
Sigaard & Skov, 2015	Empirical study via experiment	7 professionals who seek information as a job	Expectancy value theory,	The theory of expectancy-value more directly measures the effect of subjectively perceived value and perception of own capability on information-seeking behavior.
Sedera, Lokuge, Grover, Sarker, & Sarker, 2016	Empirical study via experiment	189 organization	Effectiveness, seven-point Likert scale	Digital platforms can yield innovation, only through the moderation of the enterprise system platforms.

Summary of What is Known and Unknown in Research Literature

In this section a review of literature was performed on constructs applicable to anomaly and intrusion detection. As a result, constructs relevant constructs applicable to anomaly and intrusion detection when assessing insider threat detection were identified. The presented literature review provides a summary of what is known and unknown within anomaly and intrusion detection as it pertains to identifying anomalous activities of malicious cyber insiders.

Problems within cybersecurity are different from traditional security problems, since cybersecurity issues involve human intelligence applied in order to circumvent the standard operations of a system. Solving cybersecurity issues require a different approach

(Toecker, 2014). Nations rely upon cyber systems that may inherently possess many vulnerabilities, and if exploited could be defibrillating (Harknett & Stever, 2011). It is unknown if national strategies account for technical, tactical, and operational advances in cybersecurity. Another problem within cybersecurity pertains to anomaly detection systems, which are a subset of intrusion detection systems, however, IDSs are not adequate for detecting anomalous activities (Patcha & Park, 2007). An anomaly detection system utilizing the applicable anomaly detection technique would be best suited for the detection of unknown activities (Jyothsna et al., 2011). Parsing through unknown activities to determine normal versus anomalous activities can be difficult for cyber analysts. Thus, cyber analysts struggle with analytic-focus or processing large volumes of data and providing valuable insights (Shneiderman & Plaisant, 2015). Though Shneiderman and Plaisant (2015) identified multiple focusing strategies, they were unable to determine which strategies are relevant in what situation.

Visualizations will assist cyber analysts by helping them to identify problems and work towards a viable solution (Fink et al., 2009). Using visualization interfaces in cybersecurity allows for better understanding of anomalies when interacting with data from complex cyber networks (Inibhunu et al., 2016). Though visualizations were identified as an appropriate method for anomaly identification, methods for the integration of a 'FocalPoint' and classifying an adequate display interface were to be determined. Therefore, the development of a visualization prototype that visualized already parsed data in a consumable form factor pertinent to cyber professionals better enable detection and prevention of malicious cyber insider threats.

Chapter 3

Methodology

Overview of Research Design

The study is grounded in developmental research and was conducted in three phases, as depicted in Figure 7, to develop and validate the newly designed prototype that aided in identifying anomalous activities of malicious cyber insiders. This developmental research developed a prototype, identified as a 'thing' in order to address a problem, which is considered the foundation of developmental research by Ellis and Levy (2009). The critical cyber visualization variables were identified and ranked based on SMEs' input utilizing the Delphi method. As a result, the initial version of QUICK.v™ was developed.

First, the critical cyber visualization variables are identified and ranked. Then simulated data was utilized to depict complex cyber data correlations and the top six critical cyber visualization variables to the SME participants. In order to determine the valid visualization techniques that should be applied within the developed cyber visualization prototype QUICK.v™. Next, based on the identified visualization techniques, the preliminary prototype was refined to apply the techniques identified as most valid when presenting complex data correlations and the top six critical cyber visualization variables. Subsequently, this study validated the perceived effectiveness of the developed prototype using SME analysis. To determine the perceived effectiveness of QUICK.v™, quantitative data was collected based on each SMEs rating of the

prototypes' value and their level of satisfaction. Finally, the finalized version of QUICK.v™ as well as the visualization techniques used was detailed within the findings and recommendations. Identifying and assessing the risk score is beyond the scope of this study. Risk calculation was addressed in the AI-InCyThR™ prototype (Hueca et al., 2016).

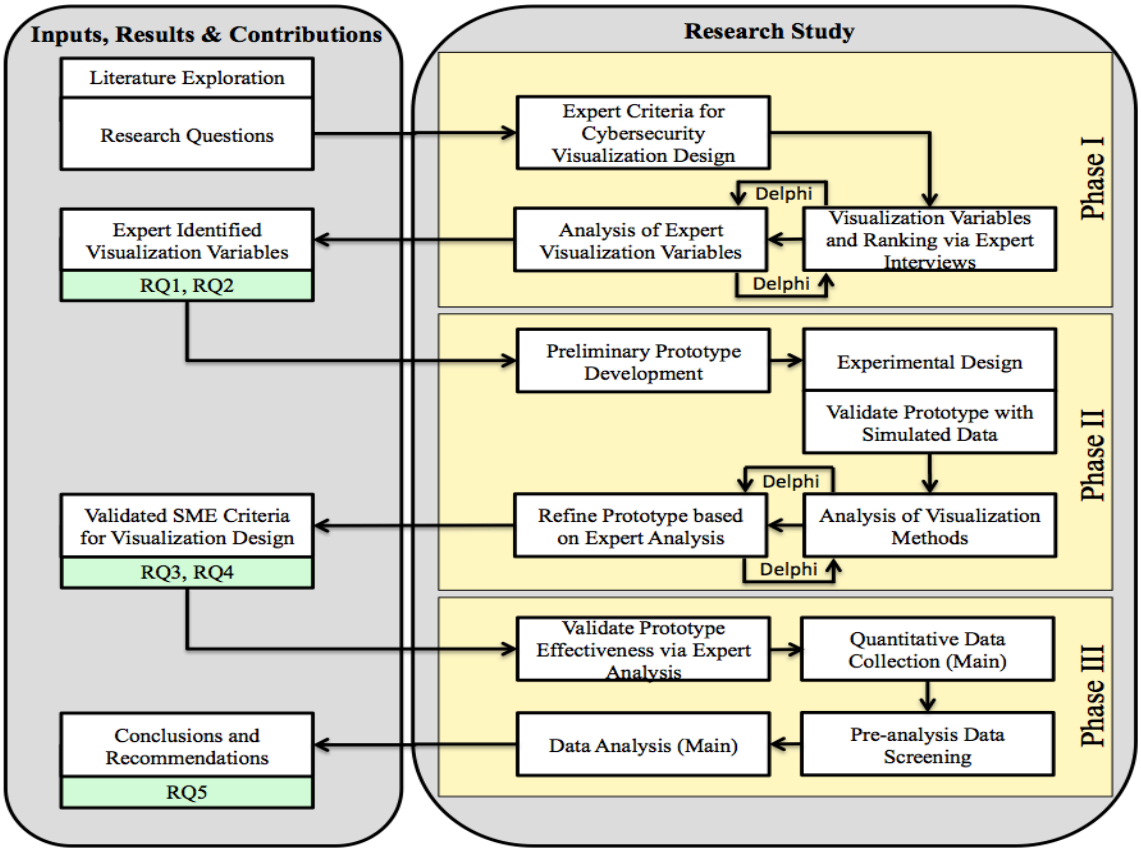


Figure 7: Research Design Process for the Development of a Cybersecurity Visualization Prototype

Phase 1

Phase one consisted of an exploratory study. This approach is generally taken when little is known about the problem at hand or no information is available on how

similar problems have been solved (Sekaran & Bougie, 2009). Legg et al. (2015) identified the insider threat topic as recently attaining attention within research and that real-world data is rarely used to address the problem. With the recent growth of big data analytics and the integration of many diverse data sources, new research solutions are needed for monitoring and analysis (Boukri & Chaoui, 2015). With the limited amount of research currently addressing the problem of detecting malicious cyber insiders, an exploratory research method is appropriate for this study. To identify the critical cyber visualization variables needed when displaying complex data correlations, extensive literature analysis was performed for the foundation of this research. With the completion of an extensive literature analysis to develop a better understanding of the research problem and formulate the research questions, each phase of this study was performed.

Phase one of this research included two parts. First, it identified the critical cyber visualization variables using SMEs that should be displayed when using applications to detect potentially malicious insider cyber threats. Second, the first phase of this study provided the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat dashboard visualization prototype to detect potentially malicious cyber insider activities.

Research question one was addressed in phase one by obtaining critical cyber visualization variables from SMEs over multiple Delphi iterations. Research question two also addressed in phase one by obtaining the rank order of the SMEs' identified critical cyber visualization variables that should be displayed. SMEs' was used to validate and rank order the critical cybersecurity visualization variables identified in literature (see

Appendix A), the developed cyber visualization prototype should incorporate that may aid in identifying potentially malicious insiders cyber threats.

SMEs were asked to review the initial list of cyber variables obtained within literature and select the most critical variables from this list. SMEs were then asked to add any additional variables they deem as critical, but are not displayed in the initial list. Next, the SMEs were asked to select their top six critical variables from the list they have created. SME responses were considered for the next round of data collection, based on prior selections. The final sets of cyber variables selected were then reviewed with the SMEs'. SMEs were asked to validate the list by adding or removing variables; once the list is finalized those variables were applied moving forward. This addressed research question one - what are SMEs' identified critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats?

After the SMEs' critical cyber visualization variables have been validated, they were asked to rank the order of the critical cyber visualization variables the developed prototype should include that may aid in identifying potentially malicious insiders cyber threats. Research question two also addressed in phase one by obtaining the ranking of the critical cyber visualization variables from SMEs over multiple Delphi iterations. The SMEs were presented the finalized list of variables and asked to rank the variables from most-to-least important. The SMEs were asked to provide a rank order for the identified variables in relation to how imperative that data is as a precursor when investigating potentially malicious cybersecurity insider threats. The highest weighted variables identified were referred to as cybersecurity 'vital signs'. The ranking was analyzed to

determine the final top six critical cyber visualization variables. This addressed research question two - what is the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat dashboard visualization prototype to detect potentially malicious cyber insider activities?

Participants in phase one consisted of at least 30 SMEs. From the group of 30 SMEs a focus group was selected to validate the final list of identified critical cyber visualization variables. Focus group participants consisted of eight to ten of the most experienced SMEs. Experience was based on number of years worked within cybersecurity, based on the SMEs demographic information. Applying the Delphi technique within a virtual lab environment, each SMEs was asked to review a list of critical cyber variables and identify the rank order of identified critical cyber visualization variables used when monitoring for potentially malicious cybersecurity insider threats.

Phase 2

Participants within phase two consisted of at least 30 SMEs. Phase two also consisted of two parts. SMEs first identified the most valid visualization techniques to present complex cyber data correlations. Second, phase two identified the visualization techniques most valid to present top six critical cyber visualization variables to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™. In phase one, the validated and ranked critical cyber visualization variables that have been collected from SMEs were applied within phase two. This consisted of SMEs' identifying the most valid presentation of complex data

correlations using the identified critical visualization variables over multiple visualization techniques.

In phase two, the preliminary prototype was developed using the ranked cybersecurity vital signs identified by the SMEs in phase one. SMEs were presented each vital sign using three different visualization techniques and asked to select their preferred visualization technique for presenting complex cyber data correlations. Once an initial set of visualization technique has been identified, the SMEs were presented options for the most utilized visualization techniques found in literature. This addressed the research question three - what SMEs' identified visualization techniques are most valid to present complex cyber data correlations relevant to the predesignated critical cyber visualization variables that are applied in the developed cyber visualization prototype QUICK.v™?

Second, phase two identified visualization techniques that are most valid when presenting the top six critical cyber visualization variables to detect potentially malicious cyber insider activities. Based on the validated and ranked critical cyber visualization variables identified in phase one, SME analysis were performed of the most valid visualization technique for presenting the top six critical cyber visualization variables. The preliminary prototype was refined to implement the validated SME criteria for the visualization design. Not all of the SME identified variables were able to be applied to QUICK.v™, the rank order from phase one was vital to determining which variables are presented using a validated visualization techniques. This addressed the RQ4 - what SMEs identified visualization techniques are most valid to present top six critical cyber visualization variables to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™?

The vital signs were presented to at least 30 cybersecurity and visualization SMEs using the most applicable visualization technique as identified in literature. Using the Delphi method the SMEs were asked to perform an analysis of the visualization techniques employed to present the complex cyber data correlations and each cybersecurity vital sign. This study controlled the visualization techniques used in the final prototype. Once SMEs have identified the critical cyber visualization variables and the most valid visualization techniques, modifications may need to be made to fit the form factor of the final prototype. Researcher interference was involved since the sample and the configuration of the prototype developed was controlled in the study.

Data collection was performed using an experiment in a contrived setting. Using a contrived setting allowed for extensive control over the experiment, as well as allow for control of external nuisance factors (Sekaran & Bougie, 2009). A lab environment was utilized for each experiment allowing quantitative and qualitative data to be captured from the participants. A lab experiment also allowed for the selection of a homogenous set of participants, or participants with similar backgrounds in cybersecurity. Selecting similar participants added to the validity of this study by allowing the measured effects to be based on a particular group (Levy & Ellis, 2011). In relation to this study the selected group was cybersecurity and visualization SMEs.

Phase 3

Participants within phase three consisted of at least 20 SMEs. Phase three entailed conducting and experimental study using SMEs' to assess the perceived effectiveness of the QUICK.v™ prototype when mitigating malicious insiders. In the third phase another virtual lab experiment was performed within a contrived setting. In this experiment the

SMEs' identified critical cyber visualization variables and visualization techniques identified within phase one and obtaining the perceived effectiveness of the developed prototype validated two. The effectiveness measure was based on the level of satisfaction and value measures obtained from the SMEs. These individual scores was used to determine the perceived effectiveness of QUICK.v™, based on the results of the data analysis performed on the quantitative data gathered from the cybersecurity SMEs. This addressed RQ5 - what is the SMEs' implied effectiveness (i.e. satisfaction and value/importance) of the QUICK.v™ prototype when mitigating potentially malicious cyber insider threats? Then, the validated QUICK.v™ prototype was presented along with recommendations and conclusions.

Instruments and Prototype Development

When visualizations are designed often times there are no explicit connections stated as to why the designer chose to utilize a particular technique Mckenna et al. (2014) addressed this problem by developing the design activity framework to directly connect each design activity with the corresponding design decision. Mckenna et al. (2015) identified the lack of developmental research utilizing cyber analyst or SME input throughout the design process. However, they struggled with obtaining direct access to cyber analysts. Using feedback from one cyber analyst qualitative coding of the current body of knowledge was applied. Additionally, Inibhunu et al. (2016) sought to increase the perceived effectiveness of cyber visualization tools by developing a system to provide adaptive level of detail in the interface. While the system was introduced the effectiveness of the system developed was not determined (Inibhunu et al., 2016).

Insider threat detection is a complex problem for cyber analysts. Walton et al. (2015) proposed a scalable visual analytics loop for continuous development of detection models. Nevertheless, there are limitations within today's cyber visualizations and additional work to advance visualizations is needed. However, input from cyber analysts is essential to develop and evaluate the perceived effectiveness, thus the resultant cyber visualization may be ready for real world application. The goal of this study is to develop and validate a cyber insider threat dashboard visualization prototype. To develop novel cyber visualizations, SME input was utilized for the development and evaluation of this research as depicted in Figure 8, a missing element among earlier works. The development process for QUICK.v™ included the following steps. Step one includes utilizing SMEs to identify the relevant cyber visualization variables. Step two consisted of identifying the rank order of the visualization variables previously identified. Step three used the identified visualization variables to determine valid visualization techniques. Step four included applying the visualization variables and the validated visualization techniques to the development of QUICK.v™. Finally, step six assessed the perceived effectiveness of QUICK.v™.

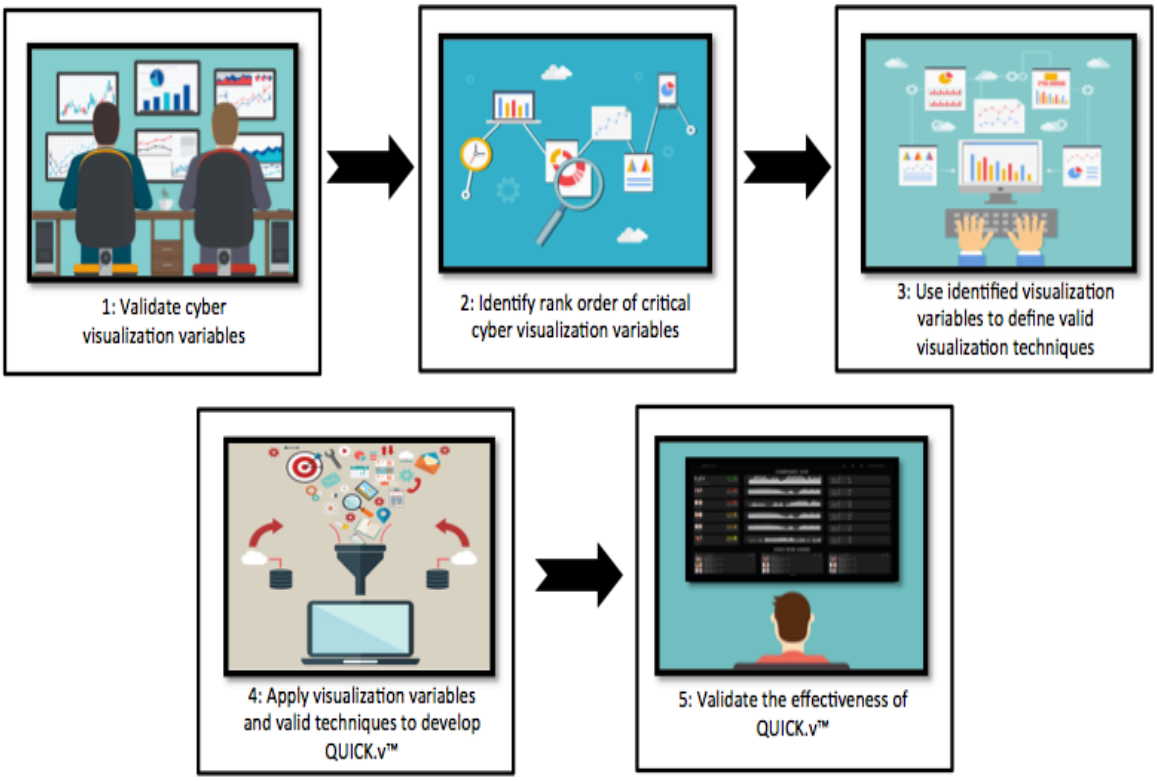


Figure 8: QUICK.v™ Development Process to Enhance Perceived Effectiveness by Assessing Satisfaction and Value

Instrument for SMEs Identification of Cyber Visualization Variables

In order to identify the cyber visualization variables SMEs were given the survey instrument presented in Appendix D. Appendix A depicts a template for how the qualitative survey was administered to the SME's. A survey refers to "gathering information about the characteristics, actions, or opinions of a large group of people, referred to as a population" (Pinsonneault & Kraemer, 1993, p. 2). For this survey the population consisted of cybersecurity analysts. By using survey research, information was gathered for measurement and understanding, thus, qualitative information was gathered to describe an aspect of the studied population, which in this case were the cyber visualization variables (Pinsonneault, & Kraemer, 1993). The survey consisted of presenting pertinent terms in relation to the scope of this research. Next, SMEs were

asked to identify data variables and rank their identified variables in relation to identifying insider threats. Variables utilized were deemed as cybersecurity vital signs that were ranked by the SME. Overlapping variables identified by the SMEs were assessed based on the denoted prioritization, then translated into linked objectives to group similar variables into cybersecurity vital signs (Keeney, 1999). The results from this survey were used to address research question one and two by identifying and ranking using SMEs the critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insiders cyber threats.

Instrument for SMEs Identification of Visualization Technique for Cyber Variables

Once the cybersecurity vital signs have been identified, the next step is to determine the appropriate visualization techniques for each vital sign. The qualitative survey in Appendix G was administered to SMEs using their criteria for visualizing each of the depicted cyber visualization variables. Appendix B depicts a template for how the qualitative survey was administered to the SME's. The final visualization techniques for each variable was updated and applied within this survey instrument once data collection and analysis utilizing Appendix D was completed. Thus, allowing for the identification of the most valid visualization techniques to present the top six critical cyber visualization variables, therefore, Appendix B was only for illustration purposes. In Appendix B segments of the preliminary prototype were presented to the SMEs depicting the cyber visualization variables they previously identified. Simulated data was used within the prototype to allow SMEs to assess and identify the data points they referenced, as well as their preferred technique of depicting the presented information. The prototype development was iterative; this involved several stages of information and user input

since initial prototypes tend to be tossed because of poorly understood requirements that are later validated based upon the preliminary prototype. However, throughout this process much experience is gained allowing for the development of a production quality prototype (Ozcan & Morrey, 1995). In this study, exclusion of the preliminary prototype may be avoided by presenting the components that make up the final prototype to the SMEs for validation. This survey consisted of presenting individual cyber visualization variables as identified by the SMEs using varying visualization techniques. The SMEs were asked to select their preferred technique for the depicted cyber visualization variable. The results from this survey, addressed research question three by identifying using SMEs the most valid visualization techniques to present complex cyber data correlations relevant to the predesignated critical cyber visualization variables.

QUICK.v™ Prototype Development

The prototype included validated visualization techniques for effective presentation of cyber visualization variables. Based on visualization techniques as identified in literature and validated by the SMEs, the techniques that are ideal for presenting the SME identified cyber related data would be identified and administered within QUICK.v™. This conceptual design is depicted in Figure 9.

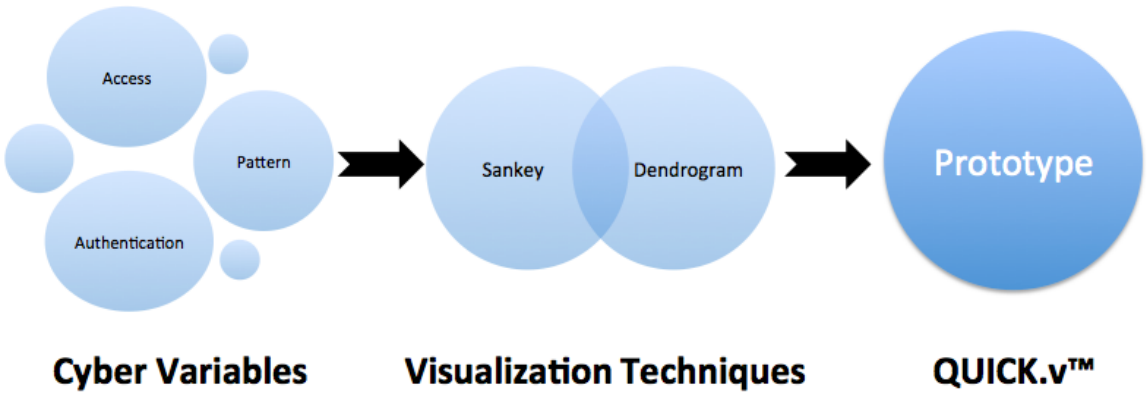


Figure 9: Conceptual Design for QUICK.v™

Once each cyber visualization variable was obtained and ranked, an ideal visualization technique was identified for the presentation of that cyber visualization variable within the display. The SMEs were able to validate or identify other manners they deem as best to depict the relevant cyber visualization variable. The validated visualization techniques were then used to develop QUICK.v™. For each identified cyber visualization variable the expert was asked to assign a weight. Based on the assigned weight of each cyber visualization variable the variable rank order was determined. Visualization techniques were identified for the pertinent cyber visualization variables. The techniques applied were then used to build the dashboard-based presentation of the QUICK.v™ prototype. Based on SME input many visualization techniques were taken into consideration for this prototype, based on literature some may include chord, sankey, dendograms, or line chart diagrams. Chou, Wang, and Ma (2016) identified sankey diagrams as being visually informative and utilized sankey diagrams to present data flows allowing for the identification of potential privacy concerns. Noel (2011) utilized dendograms to visualize cyber attack patterns since this provided an automated and mathematically sound way to present the hierarchical attack clusters. The selected visualization techniques were presented to SMEs for evaluation. SMEs were utilized to determine if the appropriate visualization technique has been selected to present the SMEs' previously identified cyber visualization variables.

Instrument for Cybersecurity Analysts' Effectiveness of the Prototype

Once the QUICK.v™ prototype has been developed based on the SME identified cyber visualization variables and visualization techniques, the perceived effectiveness of the developed prototype were measured. Using SMEs the perceived effectiveness of the

cyber visualization variables presented and the visualization techniques used were identified. Effectiveness refers to usefulness or productive use of the technology that may affect job performance and utilitarian value, which also plays a significant role in user acceptance of the technology (Levy, 2006; Dooley, Hackney, Dooley, Levy, & Parrish, 2015; Coursaris & van Osch, 2016; Levy, Hackney, & Parrish, 2017). Hong, Tai, Hwang, Kuo, and Chen (2017) utilized determinants of satisfaction and utility value within 150 questionnaires to determine effectiveness of using government e-learning systems. IS effectiveness has been difficult to evaluate. By examining the satisfaction and value of specific cyber visualization variable the perceived effectiveness of the system can be determined (Doll, Xai, & Torkzadeh, 1994). For this study, IS effectiveness was measured by obtaining user value and satisfaction (Levy, 2006). The survey in Appendix I was administered to SMEs to obtain ratings for satisfaction and value of the developed cyber visualization prototype. The survey consisted of a seven-point Likert scale assessing each cyber visualization variables and the visualization methods utilized in the final version of QUICK.v™. The survey was administered to SMEs using the online tool, survey monkey.

Expert Panel

An expert is a person qualified to address subjects from a relevant discipline (Dalkey & Helmer, 1963). To determine the relevant cyber visualization variables, SME input was obtained using the Delphi method. Based on the input obtained from the SMEs over multiple iterations of data collection, a user-centered design process was followed for the development of QUICK.v™. Capturing expert knowledge can be difficult but it is important for applying their experiences to a domain problem (Okesola, Ogunseye, &

Folorunso, 2010). Cyber SMEs tend to be extremely busy and obtaining their time can be very difficult (McKenna et al., 2015). In order to optimize the time obtained from cyber SMEs and get the most reliable consensus from the group the Delphi method was utilized (Dalkey & Helmer, 1963). Additionally, for this reason the Delphi method has been used extensively within past research (Ramim & Lichvar, 2014; Rowe & Wright, 1999; Skulmoski, Hartman, & Krahn, 2007). The Delphi method consisted of posing questions to the SMEs that are all centered on a central problem (Dalkey & Helmer, 1963).

For the purpose of this study, all questions were focused on the problem of detecting potentially malicious cyber insider threats. Prior to data collection, at least 30 SMEs in the cybersecurity or visualization industry was solicited for participation. SMEs with a background in cybersecurity was recruited from industry and consulting agencies. Since feedback was needed from the same SMEs once the prototype is completed, an experienced focus group of SMEs was sought to ensure they are still available for the prototype evaluation. SMEs were offered an incentive for their participation upon completion of phase three. Gray and Hovav (2014) denote the advantages of the Delphi method included:

- Maintaining focus on the issue
- Providing a framework for individuals to work
- Minimizing participants tendencies to agree with the leader
- Providing equal opportunity for input
- Providing documented output (p. 343)

Sun et al., (2006) as well as Ramim, Lichvar (2014) utilized Delphi techniques, as it was beneficial for obtaining consensus among SMEs on a particular topic. Thus, the Delphi

technique was used when obtaining the cyber visualization variables and rank order for the development of the visualization prototype.

Reliability and Validity

To prove that the developed instrument measures what it was intended to, the validity was pertinent to the research process (Straub, 1989). In this study data analysis needs to be conducted on three sets of data. The first point of data collection was obtained during phase one, based on the initial SME surveys the cyber visualization variables that were used to develop the prototype was determined. Then the identified cyber visualization variables were correlated to a practical visualization technique, used to depict the most simplified and understandable display of the complex data correlations. The second point of data collection was obtained within phase two, in which applicable visualization techniques were identified to display each vital sign. The visualization techniques were presented within a preliminary prototype to the SMEs for validation. The third point of data collection was once the prototype was developed to identify the perceived effectiveness of the prototype by validating the SMEs satisfaction and value pertaining to the cyber visualization variables and visualization techniques used within the developed prototype.

Reliability

The cyber visualization prototype was developed to enhance the presentation of complex data correlations when mitigating malicious insiders cyber threat. Surveys were administered to SMEs using the Delphi method. Since respondents can be inconsistent or unmotivated reliability may become an issue, which this study addressed (St. Louis,

Lubker, Yaruss, & Aliveto, 2009). SMEs were incentivized for participation, motivating them to continue throughout each phase of this study. The reliability of QUICK.v™ is determined based on the extent to which the developed product is without bias (Sekaran & Bougie, 2009). By utilizing an accepted consensus building process like the Delphi Method to obtain SME input, the reliability of this study was established (Ellis & Levy, 2010). Particular attention was also paid to variations in answers from the SMEs, since misunderstanding the questions could lead to measurement errors or irrelevant requirements (Straub, 1989).

Validity

An instrument may be invalid based on the measurement items content (Straub, 1989). Threats to internal validity include selection bias and statistical regression contamination. A threat to external validity is population validity (Sekaran & Bougie, 2009). Population validity refers to generalizing research findings from a subset of a population as applicable to the entire population. If the sample has not been randomly selected from an accessible population, the experimenter cannot generalize to a larger group of participants, the relationship between the treatment variable and the characteristics of the target population helped determine if experimenter can generalize to the target population (Bracht & Glass, 1968). If the members of an experiment are selected randomly the potential problem of statistical regression contaminating the experiment will not occur (Sekaran & Bougie, 2009). Thoroughly defining the target population and selecting the accessible population by ensuring that similar characteristics apply within both populations may reduce the potential for population validity bias (Bracht & Glass, 1968).

When conducting the experiments there is potential for selection bias to occur since the members of the participant groups needed to have a particular skill set within cybersecurity. The potential for selection bias was mitigated by randomly selecting participants from a pool of qualified participants. There was reason to assume that participants selected for the experiment are different from other employees within an organization, since particular cyber professionals were sought (Sekaran & Bougie, 2009). Having a person who is not specifically aware of the methods being tested select the participants for the experiment may control the potential for selection bias. Blackwell and Hodges (1957) states that selection bias may be eliminated by conducting the experiment in a way that the person involved in selecting the participants is not aware of the treatment methods.

The potential for experimental mortality is also a potential threat to the validity of this study. Mortality refers to the loss of participants throughout the study, which may increase the difficulty of comparing the data collected across each group of participants (Sekaran & Bougie, 2009). Since there were three phases requiring participant input, there is an increased likelihood that participants may withdraw from the study. The potential threat to participant withdrawal from the study was addressed by incentivizing participants to partake within each phase of the study. The validity and reliability of the developed prototype is pertinent to the overall study, thus, mitigation steps were taken throughout this study to reduce potential threats. Taking potential threats into consideration and using validated methods during the research design and development may reduce threats to validity and reliability (Ellis & Levy, 2010).

Prototype Perceived Effectiveness

Once QUICK.v™ has been developed the effectiveness was determined. In phase one, cyber visualization variables were obtained from at least 30 SMEs. Then in phase two the applied visualization techniques were validated using at least 30 SMEs. SME level of satisfaction and the value pertaining to the identified cyber visualization variables and applied visualization techniques were measured in phase three. Identifying the applicability of the cyber visualization variables within the final visualization was pertinent for potentially standardizing cybersecurity visualization vital signs. Rating the measure of value was deemed more beneficial than ranking characteristics, as this allows participants to denote characteristics of equal value if one did not out weight the other (Levy, 2004).

The System Usability Scale (SUS) is one of the most widely used questionnaires to measure perceived usability or user satisfaction (Lewis & Sauro, 2009). SUS allows usability practitioners to quickly and easily assess the usability of a given product or service (Bangor, Kortum, & Miller, 2008). However, SUS was designed to form a unidimensional measure of perceived usability (Borsci, Federici, Bacci, Gnaldi, & Bartolucci, 2015). Sauro (2015) used SUS to measure the convergent validity of user experiences on over 100 websites. Though some have found SUS to be a bidimensional measure, it has been focused on usability and learnability (Borsci et al., 2015). Lewis and Sauro (2017) performed an analysis of over 9,000 reports that utilized SUS, and found that the bidimensional measure of SUS was not useful. Thus, while SUS has been found to be an effective measure of usability or user satisfaction this measure primarily focuses on user satisfaction and negates the evaluation of cognitive value or level of importance when determining perceived effectiveness. Levy (2006) denoted that the complete

measurement of IS effectiveness must include measures of the causal factors or 'values' as well as the end result construct or 'user satisfaction' (p. 60). As a result this study determined perceived effectiveness using the effectiveness grid that denotes the effectiveness curves and the Value-Satisfaction Dimension Grid (Levy, 2006).

Since SUS does not measure value and satisfaction only statements from SUS was also applied within this study. Bangor et al. (2008) identified that 90% of SUS statements when utilized have been modified to compare and better understand their data since the scores from the utilized SUS statements provide a baseline score of usability. Consequentially, Bangor et al. (2008) validated the use of modified SUS statements in studies over several years, and found that modifications to SUS provide an adjective rating that correlates with a given score, and provides details of what constitutes as an acceptable SUS score. For the purpose of this study, the SUS statements depicted in Table 8 was modified and utilized in Appendix C.

Table 8

Modified SUS Statements

Original SUS Statements	Modified SUS Statements
I thought the system was easy to use	Ease of use of information depicted
I found that the various functions in this system were well integrated	Various variables were well integrated
I found the system unnecessarily complex	Complexity based on variables presented
I felt very confident using the system	Confidence quickly deciphering potential insider threats
I thought that there was too much inconsistency in this system	Consistency of visualizations presented

The perceived effectiveness of the prototype was determined using a combination of value and satisfaction in order to indicate the magnitude of the prototype effectiveness (Levy, 2006; Dooley, Hackney, Dooley, Levy, & Parrish, 2015; Levy, Hackney, &

Parrish, 2017). By allowing SMEs to evaluate the final product SMEs can denote the perceived effectiveness and potential value of QUICK.v™. The feedback and results were analyzed to determine the viability of QUICK.v™ and suggested features for future iterations.

Population and Sample

The study evaluated cyber visualization variables presented by at least 30 cyber analysts that were used to develop QUICK.v™. The unit of analysis for this study was the individual cyber analysts. The selected group of SMEs may also include cyber analytics as well as cyber forensics professionals. These SMEs were solicited specifically as the sample population for this study. A non-probability purposive judgment sampling method was applied. Professionals whom are willing to participate in the data collection were utilized to perform this study. Judgment sampling would be best applied to this research since individuals with knowledge or experience within cybersecurity and have utilized analytics tools may be in the best position to provide the information sought. This is a viable sampling method when there is only a small subset of people who possess the information needed to perform the research (Sekaran & Bougie, 2009). Cyber, cyber analytics, and cyber forensics professionals would be best suited for this study since they are more likely to be familiar with the threats malicious insiders pose, the analytics tools available and the difficulties faced when accurately detecting malicious activities within cybersecurity. A small subset of highly skilled individuals will have the identified skill set (Evans & Reeder, 2010).

Data Collection

Data collection was conducted using a series of three instruments. In phase one, data collection consisted of a qualitative survey instrument, to identify the critical cyber visualization variables (Appendix A). In phase two, data collection consisted of a qualitative survey instrument, to identify the criteria for visualization design (Appendix B). Lastly, in phase three data was collected using a quantitative survey instrument, to identify the perceived effectiveness of QUICK.v™ (Appendix C). Since data collection, analysis, and result reporting go hand in hand, accuracy of data collection was pertinent (Elo, Kaariainen, Kanste, Polkki, Utriainen, & Kyngas, 2014). To ensure accurate data is collected and that there are no missing or extreme data values pre-analysis data screening was performed (Mertler, & Vannatta 2005). This aided in ensuring accuracy when the data analysis is actually performed. Data collection was preformed on a sample of identified cybersecurity and visualization SMEs within a controlled environment. A controlled environment was utilized to allow for transcription during the data collection process.

Data Analysis

Analysis results were applied to the development of QUICK.v™. Data collection was based on both literature and input from at least 20 SMEs. This study seeks to answer research question one and research question two by using SME input. Research question three was answered using literature review and research question four was answered using a combination of both SEM input as well as literature review. Research question five was also answered based on SME provided input. Qualitative data analysis was also utilized on the dataset in phase one and two. Qualitative and quantitative data analysis was performed on this data set in phase three to obtain the rated effectiveness and

feedback in relation to each rating. Since the development of QUICK.v™ was dependent upon the cyber visualization variables identified by the SMEs. The visualization techniques used to develop QUICK.v™ are the moderating variables for the developed display. For phase one, qualitative data analysis-using coding was applied to identify the cyber visualization variables based on the SME surveys. For phase two, statistical data analysis was also conducted to determine the relevant visualization techniques by identifying the relevant means.

In phase three the perceived effectiveness of QUICK.v™ was quantified based on research analysis. The perceived effectiveness was determined using the effectiveness grid that denotes the effectiveness curves and the Value-Satisfaction Dimension Grid (Levy, 2006). Within the effectiveness grid the mean satisfaction scores are on the horizontal axis while the mean value scores are on the vertical axis (see Figure 10). Based on Levy (2006) the calculation of satisfaction was performed by determining the mean satisfaction characteristic: $\bar{S}_{a1}, \ldots \bar{S}_{a11}, \ \bar{S}_{b1}, \bar{S}_{b7}, \ \bar{S}_{c1}, \bar{S}_{c2}$. The equation that was used to compute the mean satisfaction is:

$$\bar{S}_{a1} = \left(\prod_{i=1}^{n} A_{1\_SAT} \right)_{i}^{1/n}$$

Here, A_{1_SAT} is the satisfaction score rated by SME i for cyber visualization variable A_1, and n the number of cases that the data was collected. The aggregated value score noted as $\bar{V}_{a1}, \ldots \bar{V}_{a11}, \ \bar{V}_{b1}, \ldots \bar{V}_{b7}, \ \bar{V}_{c1}, \ldots \bar{V}_{c2}$. The equation the was used to calculate the mean value is:

$$\bar{V}_{a1} = \left(\prod_{i=1}^{n} A_{1\_VAL} \right)_{i}^{1/n}$$

The A_{1_VAL} is the characteristic value score rated by SME i for cyber visualization variable A_1, and n the number of cases that the data was collected (p. 184-185). Levy (2006) developed the learners value and satisfaction (LeVIS) index in order to indicate the perceived effectiveness of e-learning systems. The LeVIS index provides an overall score of the magnitude of effectiveness for the developed prototype. Thus, this study calculated the perceived effectiveness using the LeVIS index. Based on the LeVIS index the perceived effectiveness would be determined using the formula below.

$$\left(\frac{1}{n}\right) * V_0 * S_0 \rightarrow 0 \leq LeVIS \leq 1$$

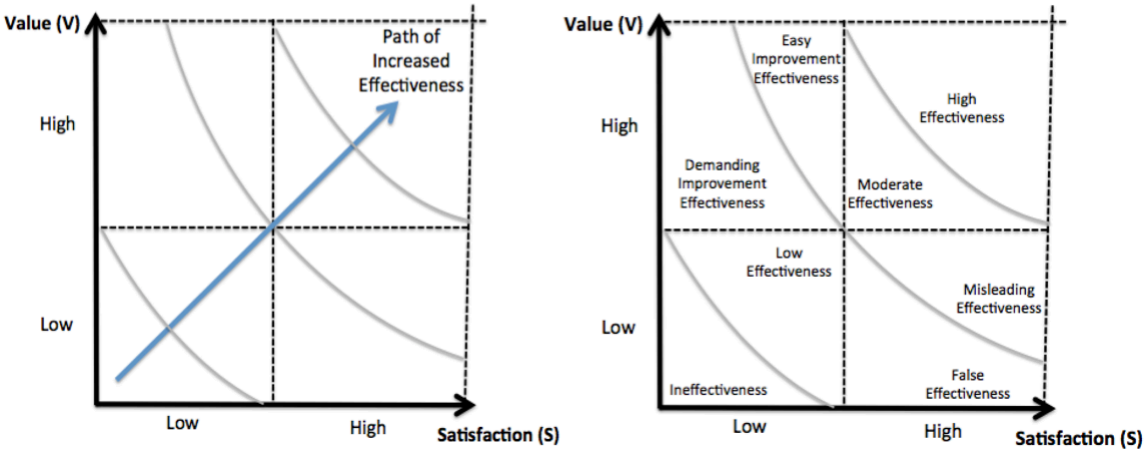


Figure 10: Effectiveness Curves & Grid
(Levy, 2006)

Pre-analysis data screening was performed on the data collected prior to fully analyzing the data. Pre-analysis data screening was performed to prevent data collection issues (Levy & Ellis, 2006). The perceived effectiveness of QUICK.v™ was dependent upon the SMEs ratings of satisfaction and importance. Multiple regression analysis may be used for data analysis. Multiple regression analysis is widely used because of its

applicability, robustness as well as ease of interpretation and may provide a baseline for evaluating empirical results (Mason, & Perreault, 1991).

An additional level of analysis was performed on QUICK.v™ to assess the usability. The modified SUS statements utilized in Appendix C was extracted for evaluation, based on the adjective rating that correlates with a given score, and correlated to an acceptable SUS score. The SMEs satisfaction rating for each item was isolated for this analysis. Each score was mapped to a mean rating and quartile. A SUS score of 70 or above was deemed as acceptable (Bangor et al., 2008). As depicted in Table 9 each adjective from Appendix C is mapped to a mean quartile. Once the rating is obtained from SMEs the SUS score and acceptability can be determined as outlined in Figure 11.

Table 9

Mapping Adjective Rating to Study Mean Quartiles

No.	***Adjective***	***M***	***Upper Bound***	***Quartile***
7	*Extremely Satisfied*	*100*	*100*	*4*
6	*Very Satisfied*	*85.58*	*87.5*	*3*
5	*Satisfied*	*72.75*	*75.0*	*2*

Mapping Adjective Rating to Study Mean Quartiles (continued)

No.	*Adjective*	*M*	*Upper Bound*	*Quartile*
4	*Neutral*	*52.01*	*55.0*	*1*
3	*Unsatisfied*	*39.17*		
2	*Very Unsatisfied*	*NA*		
1	*Extremely Unsatisfied*	*25*		

Note. NA= Not applicable

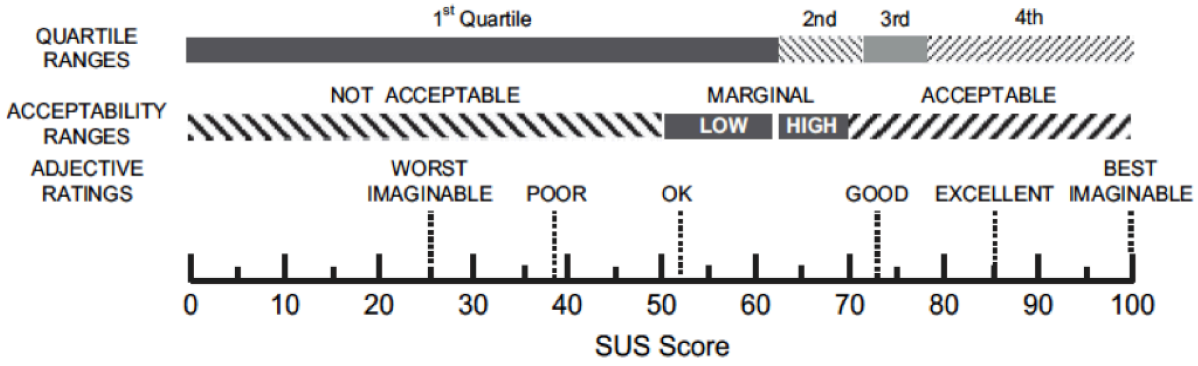


Figure 11: SUS Score by Quartile, Adjective Rating, and Acceptability (Bangor et al., 2008)

Resources

This study may require Institutional Review Board (IRB) approval since human participants was involved when testing the developed prototypes. Cybersecurity and visualization SMEs need to be obtained for applying the Delphi method for SME input. An HTML5 developer was utilized to develop the final prototype which required a website and hosting to allow easy access from the web. In addition this study also required obtaining parsed feeds of simulated data sets, this was housed in an AWS database. Access to a mobile device or computer was needed for presenting the developed visualization. Finally, ten $10 gift cards were needed to provide to the SMEs as a reward for their contributions to this study.

Summary

In this chapter an overview of the research methodology utilized for this study was provided. This study used a mixed methods approach, thus, quantitative and

qualitative data was incorporated to develop, validate, and test the perceived effectiveness of a newly designed prototype that aided in identifying anomalous activities of malicious cyber insiders. The research questions that was addressed are:

1. What are SMEs' identified critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats?
2. What is the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat dashboard visualization prototype to detect potentially malicious cyber insider activities?
3. What SMEs' identified visualization techniques are most valid to present complex cyber data *correlations* relevant to the predesignated critical cyber visualization variables that are applied within the developed cyber visualization prototype QUICK.v™?
4. What SMEs' identified visualization techniques are most valid to present *top six critical cyber visualization variables* to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™?
5. What is the SMEs' perceived effectiveness (i.e. satisfaction & value/importance) of the QUICK.v™ prototype when mitigating potentially malicious cyber insider threats?

To address the research questions, the research methodology was applied over three phases. In phase one, an exploratory study was performed to identify the criteria

needed for displaying complex data correlations and visualization techniques extensive literature analysis is required for the foundation of the research. In phase two, an experimental study was performed to identify the criteria or user requirements for the visualization design. In phase three, qualitative data analysis was performed on data collected via an experiment to determine the perceived effectiveness or usability of the developed prototype. Obtaining a SUS score also provided a very useful metric for the overall prototype usability (Bangor et al., 2008). Survey instruments were utilized in each phase to obtain applicable data, which was applied to the developed prototype. A survey instrument was developed to identify the cyber visualization variables SMEs. A qualitative survey instrument was developed to identify the visualization techniques for the identified cyber visualization variables obtained from instrument one.

A prototype was developed based on the results of the data obtained from both instruments. Lastly, an instrument was utilized to determine the perceived effectiveness of the develop prototype. The population samples consisted of cybersecurity and visualization SMEs. The chapter concluded with an overview of the resources that were utilized in completing this experimental study.

Chapter 4

Results

Overview

Upon completion of the data collection process the methods of statistical analysis and the data analysis process utilized are included in this chapter. In phase one, the results are presented based on the data collection using SMEs and applying the Delphi Method to identify and rank the most critical cyber visualization variables. Then, the results from phase one are presented. In phase two, data collection is performed using SMEs by applying the Delphi Method to validate the identified visualization techniques and the techniques most valid for presenting complex cyber data. In phase three, the data analysis and process used to determine the perceived effectiveness of QUICK.v™ was completed. Based on data collection performed using SMEs, the identified value and satisfaction for the developed prototype QUICK.v™ was attained. Presented at the end of this chapter is the summary of the results for all three phases.

Phase One - Expert Panel

Initial data collection compromised of data collected from cybersecurity SMEs. For phase one of this study, using the Delphi Method data collection was conducted early January 2018 to late February 2018. The following sections present the data collection process for phase one.

Phase One – Data Collection

For phase one of this study, the goal for the SMEs was to identify the critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats. Prior to the initial survey a pilot was performed with five SMEs to verify the reliability of the data. Pilot studies are important for trying out the research instrument as this could identify points where the proposed instrument may be complicated or fail (Teijlingen & Hundley, 2001). The survey was then refined and the final instrument used for phase one is presented in Appendix E. The SMEs consisted of 300 cybersecurity professionals. These individuals were sourced from LinkedIn social network. They included individuals in academia, public, and private sector companies. Individuals in this group were located in the U.S., Europe, and India. These individuals were selected as described in chapter three. An email presented in Appendix D containing a link to a Web-based survey tool was used to record the responses of the SMEs. A total of 42 SMEs completed the phase one survey. Upon completing the survey another round of data collection for the identified variables was performed with eight SMEs, additional qualitative data was captured in relation to each variable and their selections.

Phase One – Pre-Analysis Data Screening

Prior to data analysis pre-analysis data screening was performed on the data collected from the SMEs. As noted by Mertler and Vannatta (2005), to ensure accurate data is collected and that there are no missing or extreme data values pre-analysis data screening needs to be performed. Levy and Ellis (2006) also denoted pre-analysis data screening needs to be performed to prevent data collection issues. SME responses were

collected using the Web-based tool SurveyMonkey®, which technically ensure completeness by impeding impartial survey submissions. As a result, none of the surveys submitted were excluded. Upon performing pre-analysis data screening no outliers were identified or excluded, thus, all 42 responses collected were complete and included in the data analysis.

Phase One - Expert Panel Characteristics

For phase one, the SMEs were solicited from LinkedIn professional contacts. A survey was then distributed via email to collect responses. A total of 42 SMEs responded by completing survey instrument one. This survey consisted of questions concerning the identification of potentially malicious cyber insiders by identifying relevant cyber visualization variables. Respondents were then asked to identify relevant demographic information. Demographic information requested from the SMEs included gender and age. To identify the expertise level of the SMEs they were asked to identify the number of professional certifications they currently hold between zero and five or more certifications. The level of education of the SMEs was also identified. SMEs were to identify themselves as holding a 2-year college (Associates degree), a 4-year college (Bachelors degree), a Graduate degree, or a Doctorate.

SMEs were also asked to identify their experience within their current roles based on the number of years they have worked within their current organizations. They were able to select from under one year, one to five years, six to 10 years, 11 to 15 years, 16 to 20 years, and over 30 years. SMEs were to identify themselves as working within academia, federal government, private sector, sate government, or other. SMEs were asked to identify their job function within cybersecurity. Finally, the SMEs were asked to

describe their current employer by identifying which role described their current employer. Table 10 represents the complete SME demographic distributions based on their responses.

Table 10

Demographic Distribution of the SMEs (N= 42)

		Frequency	Percentage (%)
Gender			
	Female	4	9.5%
	Male	38	90.5%
Age Category			
	20-29	11	26.2%
	30-39	13	31.0%
	40-49	13	31.0%
	50-59	5	11.9%
Certification			
	0	17	40.5%
	1	9	21.4%
	2	3	7.1%
	3	4	9.5%
	4	3	7.1%
	5 or more	6	14.3%
Education			
	2-year college (Associates degree)	1	2.4%
	4-year college (Bachelors degree)	18	42.9%
	Graduate degree	19	45.2%
	Doctorate	4	9.5%
Experience			
	Under 1 year	7	20.0%
	1 - 5 years	23	65.7%
	6 - 10 years	6	17.1%
	11 - 15 years	2	5.7%
	16 - 20 years	3	8.6%
	Over 30 years	1	2.9%
Employer			
	Academia	5	11.9%
	Federal government employee	4	9.5%
	Private sector company	29	69.0%
	State government employee	1	2.4%
	Other	3	7.1%

Demographic Distribution of the SMEs (continued)

		Frequency	Percentage (%)
Job Function			
	Cybersecurity Administrator	1	2.4%
	Cybersecurity Analyst	1	2.4%
	Cybersecurity Architect	2	4.8%
	Cybersecurity Consultant	15	35.7%
	Cybersecurity Engineer	4	9.5%
	Information Assurance Engineer	1	2.4%
	Information Security Analyst	1	2.4%
	Information Security Manager	4	9.5%
	Network Security Engineer	1	2.4%
	Other	12	28.6%

Phase One – Data Analysis

In phase one, the data collected was exported to Microsoft Excel for initial analysis. Answers to each survey question were parsed to identify the count of each variable selected by the SMEs. In survey one the variables within each individual category were first assessed by the SMEs. To address RQ1, what are SMEs' identified critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats? SMEs were asked to select the relevant analytic variable within each identified category that they deem to be most important when trying to identify potentially malicious insider threats. The categories of analytic variables individually assessed include: System, Social, Health, Human Resources, Financial, Security, and Criminal.

Since the SMEs were asked to select at most two analytic variables for each category, the variables were weighted based on their selection for analysis. If the SME did not deem a variable as critical, they did not have to select a variable within that

category. Thus, some variables were not selected and had a count of zero. The count or number of responses pertaining to the particular variable represents the total instances that an analytic variable was selected as a response for variable one and variable two within each category. If the variable was selected as variable one or variable two, this selection was then weighted to obtain the weighted average ranking for that variable. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question. The weighted average ranking was then converted to a percentage to represent the final ranking of the applicable variable within its category. For instance, in Table 11 for the analytic variable 'access inconsistent with user class', the count of SME selection for variable one was eight, while the count for variable two was also eight, resulting in a total 'Count' of sixteen. The eight selections for variable one and two were subsequently weighted, with variable one being given a higher weight than variable two. The weighted average ranking for the variable was then determined by using the weight of ranked position (w), multiplied by the response count (x), as depicted in the formula below.

$$\frac{x_1 w_1 + x_2 w_2}{Total}$$

Once the weighted average ranking for all variables within the category was determined, the applicable percentage for that raking in relation to all variables were determined. This was performed for each variable within the category. Subsequently, the same steps were performed for each individual analytic variable category.

In this study, system analytic variables refer to system generated events based on event logs. System event logs generally portray system, user, and network activity. These logs are generated by the server operating system, firewall, or other applications within

an organizations environment (Girardin & Brodbeck, 1998). Audit logs contain a detailed trace of an operating system (Yoo, Jo, Kim, & Seo, 2018). From these logs observations relevant to user behavior can be made. The system analytic variables below reflect types of observations a cybersecurity analyst can make based on the event logs. Table 11 represents the final weight allocation for the system analytic variables based on SME responses.

Table 11

Weight Allocations for the Two Relevant System Analytic Variables Selected (N= 42)

System - Analytic Variables	Count	%
Access Inconsistent With User Class	16	19%
Changes in Data Access Patterns	12	16%
Privilege Change	10	14%
Authentication and Authorization Failure	9	14%
Data Exfiltration	7	12%
Unauthorized Data Access Methods	13	8%
Audit Log Modification	6	7%
Knowledge Access	2	4%
Network Patterns Inconsistent with User Class	4	3%
Improper Command Usage	1	2%
Changes in Network Patterns	1	1%
Erroneous Defensive Posture Changes	0	0%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, facility analytic variables refer to event logs generated in relation to aspects of physical access. The times that a user enters or leaves a facility. This data is generally logged and monitored. Using logs that show physical movement via the system access control systems can aid in the detection of malicious insiders (Sanders, 2017). Access times and locations are vital to protecting cybersecurity systems within an organization (Denning & MacDoran, 1996). The facility analytic variables below reflect types of observations a cybersecurity analyst can make based on these event logs. Table

12 represents the final weight allocation for the facility analytic variables based on SME responses.

Table 12

Weight Allocations for the Two Relevant Facility Analytic Variables Selected (N= 42)

Facility - Analytic Variables	Count	%
Time of Access Pattern Changes	39	71%
Locality of Access Pattern Changes	41	29%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the business capability analytic variables refer to aspects of an organizations business model that define their business capabilities. Business capabilities are generally focused on people, process, and technology within the organization Rosemann & vom Brocke, 2015). By integrating, monitoring and analyzing a vast amount of dispersed event logs organizations can monitor and analyze the performance of their business processes (Vera-Baquero, Colomo Palacios, Stantchev, & Molloy, 2015). The business capability analytic variables below reflect types of observations a cybersecurity analyst can make based on relevant business capability events. Table 13 represents the final weight allocation for the business capability analytic variables based on SME responses.

Table 13

Weight Allocations for the Two Relevant Business Capability Analytic Variables Selected (N= 42)

Business Capability - Analytic Variables	Count	%
Malware Deployment	16	35%
Failure Correlation	11	24%
Attribution of Disclosure	5	18%
Analysis of Competitor	4	12%
Retrieval	8	8%
Analysis of Public Media	3	3%
Deletion or Modification of Data or Infrastructure	0	0%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the social analytic variables refer to aspects of a users behavior or relevant human factors. Negative attitudes toward business activities and impulsivity can be correlated with risky cybersecurity behaviors (Hadlington, 2017). Monitoring human factors in cybersecurity is important, if ignored these factors can place an organization at risk. Since social analytic variables does not generally have data available within event logs or SIEM solutions. This data may be identified using dictionary languages with theme-specific dictionaries that contain a list of words that have been validated to be associated with constructs. These constructs may include wellbeing, engagement, positive and negative emotion, power, etc. In the future this data may be collected from employee email, social media, or instant messenger content (Shami, Muller, Pal, Masli, & Geyer, 2015). The social analytic variables below reflect types of observations a cybersecurity analyst can make based on relevant social events. Table 14 represents the final weight allocation for the social analytic variables based on SME responses.

Table 14

Weight Allocations for the Two Relevant Social Analytic Variables Selected (N= 42)

Social - Analytic Variables	Count	%
Unauthorized or Inappropriate Associations	23	37%
Personal Inflexibility	7	12%
Workplace Satisfaction	11	10%
Disregard	6	10%
Unusual Contacts	10	8%
Unusual Business Travel	7	8%
Unusual Personal Travel	3	7%
Withdrawal	8	5%
Workplace Events	4	3%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the health analytic variables refer to aspects of a users behavior or relevant human factors. Human factors contribute to cybersecurity vulnerabilities and risks (Hadlington, 2017). Instability and impulsivity can impact a users tendency toward malicious activities (King, Henshel, Flora, Cains, Hoffman, & Sample, 2018). The health analytic variables below reflect types of observations a cybersecurity analyst can make based on relevant health events. Table 15 represents the final weight allocation for the health analytic variables based on SME responses.

Table 15

Weight Allocations for the Two Relevant Health Analytic Variables Selected (N= 42)

Health - Analytic Variables	Count	%
Mental instability	40	62%
Impulse Control	43	38%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the human resource analytic variables refer to events obtained via human resources (HR). Generally, HR representatives track and monitor complaints or review data pertaining to a user. Human characteristics can be correlated with

cybersecurity behaviors, this data can be utilized to identify cybersecurity events (Gratian, Bandi, Cukier, Dykstra, & Ginther, 2018). The HR analytic variables below reflect types of observations that can be made based on relevant HR events. Table 16 represents the final weight allocation for the health analytic variables based on SME responses.

Table 16

Weight Allocations for the Two Relevant Human Resource Analytic Variables Selected (N= 42)

Human Resources - Analytic Variables	Count	%
Complaints Against the User	37	45%
Major Life Event	18	31%
Negative Reviews	27	24%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the financial analytic variables refer to events related to a users financial means. A users financial means or changes to those means can affect users cybersecurity related activities. With changes in means a user can generally be tied to fraudulent insider activities (Westerlund, Craigen, Bailetti, & Agwae, 2018). The financial analytic variables below reflect types of observations that can be made based on relevant financial events. Table 17 represents the final weight allocation for the financial analytic variables based on SME responses.

Table 17

Weight Allocations for the Two Relevant Financial Analytic Variables Selected (N= 42)

Financial - Analytic Variables	Count	%
Observed Change in Means Relative to Peers	25	36%
Observed Temporal Change in Means	26	33%
Financial Reporting	28	31%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the security analytic variables refer to user activities that trigger security related events. Security analytic variables are generally identified using event logs. An organization may identify specific triggers for what constitutes a security event. For the purpose of this study examples of security events may include: software installation, managing system services, or successful and failed login attempts (Malec, Piwowar, Kozakiewicz, & Lasota, 2015). The security analytic variables below reflect types of observations that can be made based on relevant security events. Table 18 represents the final weight allocation for the security analytic variables based on SME responses.

Table 18

Weight Allocations for the Two Relevant Security Analytic Variables Selected (N= 42)

Security - Analytic Variables	Count	%
Unauthorized or Inappropriate Use of Tools	33	58%
Duration and Regularity of Security Events	23	27%
Change in Violation Patterns	23	15%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

In this study, the criminal analytic variables refer to events related to a users criminal activity. Criminal activities may contribute to a users cybersecurity related

activities. (Chang, Zhong, & Grabosky, 2018). The criminal analytic variables below reflect types of observations that can be made based on relevant criminal events. Table 19 represents the final weight allocation for the criminal analytic variables based on SME responses.

Table 19

Weight Allocations for the Two Relevant Criminal Analytic Variables Selected (N= 42)

Criminal - Analytic Variables	Count	%
Recent Increase in Criminal Events	28	40%
Violence Outside Workplace	25	25%
Wage Garnishments	15	25%
Restraining Orders	9	10%

Note. The total count or maximum number of potential responses was 84, since each of the 42 participants had the option of selecting two variables per question.

Critical Cyber Visualization Variables Rank Order

Research question two asked: what is the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat dashboard visualization prototype to detect potentially malicious cyber insider activities? To address it, in survey instrument one, SMEs were presented with all 45 analytic variables. They were then asked to think of issues related to the insider threat detection and list only the top five variables that are the most important to them when identifying malicious insider threats. From the SMEs weighted rankings the identified top six critical cyber visualization variables were: workplace satisfaction, change in violation patterns, audit log modification, change in data access patterns, data exfiltration, and privilege change. Table 20 presents the resulting rank order for all critical cyber visualization variables based on SME responses.

Table 20

Final Weighted Rankings for all Analytic Variables Selected as the Top Five Most Critical When Identifying Potentially Malicious Insider Threats

Analytic Variables	Weighted Mean	Final Rank
Social-Workplace Satisfaction	4.00	1
Security-Change in Violation Patterns	4.00	1
System-Audit Log Modification	3.82	3
System-Changes in Data Access Patterns	3.80	4
System-Data Exfiltration	3.70	5
System-Privilege Change	3.69	6
Business Capabilities-Malware Deployment	3.60	7
Human Resources-Negative Reviews	3.60	7
System-Authentication and Authorization Failure	3.45	9
System-Unauthorized Data Access Methods	3.36	10
Facility-Time of Access Pattern Changes	3.29	11
Facility-Locality of Access Pattern Changes	3.00	12
Criminal-Recent Increase in Criminal Events	3.00	12
System-Access Inconsistent With User Class	2.78	14
Criminal-Violence Outside Workplace	2.67	15
System-Changes in Network Patterns	2.50	16
Criminal-Wage Garnishments	2.40	17
Human Resources-Complaints Against the User	2.25	18
System-Knowledge Access	2.20	19
Social-Unauthorized or Inappropriate Associations	2.17	20
Business Capabilities-Deletion or Modification of Data or Infrastructure	2.10	21
Business Capabilities-Failure Correlation	2.00	22
Social-Disregard	2.00	22
Social-Workplace Events	2.00	22
Human Resources-Major Life Event	2.00	22
Security-Duration and Regularity of Security Events	2.00	22
System-Network Patterns Inconsistent with User Class	1.50	27
Security-Unauthorized or Inappropriate Use of Tools	1.50	27
Business Capabilities-Analysis of Public Media	1.00	29
Social-Unusual Business Travel	1.00	29
Social-Unusual Personal Travel	1.00	29
Social-Withdrawal	1.00	29
Criminal-Restraining Orders	1.00	29

Analytic Variable Final Weighted Rankings (continued)

Analytic Variables	Weighted Mean	Final Rank
System-Improper Command Usage	0.00	34
System-Erroneous Defensive Posture Changes	0.00	34
Business Capabilities-Attribution of Disclosure	0.00	34
Business Capabilities-Analysis of Competitor	0.00	34
Business Capabilities-Retrieval	0.00	34
Social-Personal Inflexibility	0.00	34
Social-Unusual Contacts	0.00	34
Social-Mental instability	0.00	34
Social-Impulse Control	0.00	34
Human Resources-Observed Change in Means Relative to Peers	0.00	34
Human Resources-Observed Temporal Change in Means	0.00	34
Human Resources-Financial Reporting	0.00	34

Phase One - Comments

No exclusions were identified within the 42 completed surveys. In survey instrument one, SMEs where asked to comment on the positives and negatives associated with their identified top five critical cyber visualization variables. Commenting was not required for these questions. The responses received were relatively consistent and general in nature. For the positives associated with the top five critical cyber visualization variables similar comments suggested the variables identified would: be a good indicator of a potential problem, allow for a quicker way to pin-point an issue, and assist with quicker resolutions to detected serious violations. The negatives associated with the top six critical cyber visualization variables identified by the SMEs suggested: identified violations based on the variables should be considered in context, false positives may persist as an issue, and that some data may be difficult for an employer to obtain. In addition to the comments within the survey the focus group was contacted for additional

validation of the data collected. Comments from the focus group aligned with the comments obtained within survey instrument one as a result no further additions or collection was needed.

Phase Two - Expert Panel

Initial data collection compromised of data collected from cybersecurity and visualization SMEs. For phase two of this study, using the Delphi Method data collection was conducted mid-March 2018 to late March 2018. The following sections present the data collection process for phase two.

Phase Two – Data Collection

For phase two of this study, the goal for the SMEs was to identify the visualization techniques that are most valid to present the top five critical cyber visualization variables. Also, to identify the visualization techniques that are most valid to present complex cyber data correlations. Prior to the initial survey, a pilot was performed with three SMEs to verify the survey's ability to collect appropriate data. The survey was then refined and the final instrument used for phase two is presented in Appendix G. The SMEs consisted of 80 cybersecurity and visualization professionals, these individuals were sourced from the SME demographic information collected in phase one. SMEs with higher levels of expertise were invited to participate. Additional SMEs were added to the list from the initial 300 participants identified for phase one, based on certification and job function data on their LinkedIn profiles. The SMEs for phase two included individuals in academia, public, and private sector companies. Individuals in this group were located primarily in the U.S. An invitation was sent to each participant to complete the Google Forms® presented in Appendix F. A total of 31 SMEs

completed the survey in this round. Upon completing the survey another round of data collection for the identified variables was performed with six SMEs from the focus group, additional qualitative data was captured in relation to each variable and their selections.

Phase Two – Pre-Analysis Data Screening

Prior to data analysis pre-analysis data screening was performed on the data collected from the SMEs. SME responses were collected using the Web-based tool Google Forms®, this tool allowed for technical restrictions to form submissions without completing all questions. This ensured completeness by impeding impartial survey submissions. As a result and given no missing data existed, none of the surveys submitted were excluded. Upon performing pre-analysis data screening no outliers were identified or excluded. The 31 responses collected were all included for data analysis.

Phase Two - Expert Panel Characteristics

For phase two, the SMEs were solicited from the researches LinkedIn professional contacts. A survey was then distributed via email to collect responses. A total of 31 SMEs responded by completing survey instrument two. This survey consisted of questions concerning the identification of visualization techniques to present top five critical cyber visualization variables and to present complex cyber data correlations. Respondents were asked to identify relevant demographic information. Demographic information requested from the SMEs included gender and age. To identify the expertise level of the SMEs they were asked to identify the number of professional certifications they currently hold between zero to five or more certifications. The level of education of the SMEs was also identified. SMEs were to identify themselves as holding a 2-year

college (Associates degree), a 4-year college (Bachelors degree), a Graduate degree (MA/MS), or a Doctorate.

Additionally, SMEs were asked to identify their experience within their current roles based on the number of years they have worked within their current organizations. They were able to select from under one year, one to five years, six to 10 years, 11 to 15 years, 16 to 20 years, and over 30 years. The SMEs were asked to describe their current employer by identifying which role described their current employer. They were to identify themselves as working within academia, federal government, private sector, sate government, or other. SMEs were asked to identify their job function within cybersecurity. Table 21 represents the demographic distribution of their responses.

Table 21

Demographic Distribution of the SMEs (N= 31)

		Frequency	%
Gender			
	Female	6	19.4%
	Male	25	80.6%
Age Category			
	20-29	12	38.7%
	30-39	10	32.3%
	40-49	5	16.1%
	50-59	4	12.9%
Certification			
	0	9	29.0%
	1	8	25.8%
	2	6	19.4%
	3	1	3.2%
	4	1	3.2%
	5 or more	6	19.4%
Education			
	2-year college (Associates degree)	1	3.2%
	4-year college (Bachelors degree)	18	58.1%
	Graduate degree (MA/MS)	12	38.7%
	Doctorate	0	0.0%

Demographic Distribution of the SMEs (continued)

		Frequency	%
Experience			
	Under 1 year	10	47.6%
	1 - 5 years	16	76.2%
	6 - 10 years	1	4.8%
	11 - 15 years	3	14.3%
	16 - 20 years	0	0.0%
	Over 30 years	1	4.8%
Employer			
	Academia	5	16.1%
	Federal government employee	2	6.5%
	Private sector company	1	3.2%
	State government employee	1	3.2%
	Other	22	71.0%
Job Function			
	Cybersecurity Administrator	1	3.2%
	Cybersecurity Analyst	0	0.0%
	Cybersecurity Architect	2	6.5%
	Cybersecurity Consultant	14	45.2%
	Cybersecurity Engineer	6	19.4%
	Information Assurance Engineer	0	0.0%
	Information Security Analyst	0	0.0%
	Information Security Manager	4	12.9%
	Network Security Engineer	0	0.0%
	Other	4	12.9%

Phase Two – Data Analysis

In phase two the data collected was converted to Excel for initial analysis, answers to each survey question were parsed to identify the count of each variable selected by the SMEs. The visualization technique for each of the SMEs' identified critical cyber visualization variable were first assessed. RQ3 asked: what SMEs' identified visualization techniques are most valid to present complex cyber data correlations relevant to the pre-designated critical cyber visualization variables that are applied within the developed cyber visualization prototype QUICK.v™? To address

RQ3, SMEs were presented with three options of visualization techniques identified for presenting complex data correlations. They were then asked to select the visualization technique from the three options presented that are most relevant to displaying two or more data variables. RQ4 four asked: what SMEs' identified visualization techniques are most valid to present top six critical cyber visualization variables to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™? To address RQ4, SMEs were presented with three options of visualization techniques identified for presenting each of the six critical cyber visualization variables. Based on the data analysis results from phase one, since six critical cyber visualization variables were identified, survey two was modified to add six critical cyber visualization variables. SMEs were asked to select the visualization technique from the three options presented that are most relevant to displaying data related to each of the six critical cyber visualization variables.

Since the SMEs were asked to select one option for each visualization technique presented the responses were not weighted prior to determining the final allocations. All questions were also marked are requiring a response, as a result SMEs had to select a visualization technique for each of the critical cyber visualization variables. Table 22 through Table 27 represent the final rankings for the present top six critical cyber visualization variables. Based on the count (n), the average for each visualization technique was identified. This was performed on all data collected for each of the six critical cyber visualization variables, allowing for the identification of which visualization techniques are most valid to present each of the top six critical cyber visualization variables.

Table 22

Visualization Technique Rankings (N= 31)

		N	%
Workplace Satisfaction	Line Graph	16	52%
	Bar Graph	13	42%
	Calendar View	2	6%
Change in Violation Pattern	Area Chart	18	58%
	Radar Plot	9	29%
	Streamgraph	4	13%
Audit Log Modification	Line Graph	18	58%
	Fisheye Distortion	8	26%
	Bar Graph	5	16%
Change in Data Access Pattern	Stacked Column Graph	18	58%
	Stacked Bar Graph	12	39%
	Streamgraph	1	3%
Data Exfiltration	Line Graph	18	58%
	Column Graph	9	29%
	Fisheye Distortion	4	13%
Privilege Change	Line Graph	16	52%
	Stacked Bars	14	45%
	Stacked Columns	1	3%
Complex Data Correlations	Parallel Coordinates	15	48%
	Chord Diagram	14	45%
	Hierarchical Bundling	2	6%

Phase Two – Comments

No exclusions were identified within the 31 completed surveys. In survey instrument two, SMEs were not asked for comments. Since the identified list of SMEs

consisted of 80 individuals, using the Delphi method the initial set of surveys were sent to all 80 SMEs to complete on their own. Subsequently, survey instrument two was again administered to eight SMEs of the focus group using a contrived setting allowed for extensive control over the experiment. A virtual lab environment was used via Google Hangout to obtain qualitative data from these SMEs. Consistent responses were received. SMEs denoted though they were intrigued by the more creative visualization techniques presented, they veered towards selecting the more standard methods that they are used in other contexts and could easily make sense of. Another recurrent comment was, that SMEs are often presented with dashboards that make no sense to them and found this to be a good way to start standardization. All other comments were generally around data within the visualizations, to which SMEs' were then asked, for this survey to bring their focus back on the visualization technique itself and not the data presented.

Phase Three – Expert Panel

Initial data collection compromised of data collected from cybersecurity and visualization SMEs. For phase three of this study, data collection was conducted mid-April 2018 to late April 2018. The following sections present the data collection process for phase three.

Phase Three – Data Collection

In phase three of this study, the goal for the SMEs was to identify the perceived effectiveness (i.e. satisfaction & value/importance) of the developed prototype QUICK.v™. Prior to the initial survey a pilot was performed with five SMEs to verify the survey's ability to collect appropriate data. The survey was then refined and the final

instrument used for phase three is presented in Appendix I. The SMEs in this phase consisted of 300 cybersecurity and visualization professionals. These individuals were sourced from LinkedIn network. They included individuals in academia, public, and private sector companies. Individuals in this group were located primarily in the U.S. and India. An invitation was sent to each participant to complete the Google Forms® presented in Appendix H. A total of 26 SMEs completed the survey.

Phase Three – Pre-Analysis Data Screening

Prior to data analysis pre-analysis data screening was performed on the data collected from the SMEs. SME responses were collected using the Web-based tool Google Forms®, this tool allowed for technical restrictions to form submissions without completing all questions. This ensured completeness by impeding impartial survey submissions. Upon performing pre-analysis data screening, an outlier was identified and excluded from further analysis. No additional responses were excluded, thus from the 26 responses collected, 25 were included for data analysis.

Phase Three - Expert Panel Characteristics

For phase three, the SMEs were solicited from LinkedIn professional contacts. A survey was then distributed via email to collect responses. A total of 26 SMEs responded by completing survey instrument one was usable due to one outlier. This survey consisted of questions to identify the level of satisfaction and the value for each of the top six critical cyber visualization variables and presentation technique. Respondents were asked to identify relevant demographic information. Demographic information requested from the SMEs included gender and age. To identify the expertise level of the SMEs they were asked to identify the number of professional certifications they currently hold between

zero to five or more certifications. The level of education of the SMEs was also identified. SMEs were to identify themselves as holding a 2-year college (Associates degree), a 4-year college (Bachelors degree), a Graduate degree (MA/MS), or a Doctorate.

Additionally, SMEs were asked to identify their experience within their current roles based on the number of years they have worked within their current organizations. They were able to select from under one year, one to five years, six to 10 years, 11 to 15 years, 16 to 20 years, and over 30 years. The SMEs were asked to describe their current employer by identifying which role described their current employer. They were to identify themselves as working within academia, federal government, private sector, sate government, or other. SMEs were asked to identify their job function within cybersecurity. Table 23 represents the demographic distribution of their responses.

Table 23

Demographic Distribution of the SMEs (N= 25)

		Frequency	%
Gender			
	Female	4	16.0%
	Male	21	84.0%
Age Category			
	20-29	7	28.0%
	30-39	11	44.0%
	40-49	4	16.0%
	50-59	3	12.0%
Certification			
	0	9	36.0%
	1	8	32.0%
	2	3	12.0%
	3	2	8.0%
	4	1	4.0%
	5 or more	2	8.0%

Demographic Distribution of the SMEs (continued)

		Frequency	%
Education			
	2-year college (Associates degree)	2	8.0%
	4-year college (Bachelors degree)	10	40.0%
	Graduate degree (MA/MS)	11	44.0%
	Doctorate	2	8.0%
Experience			
	Under 1 year	3	13.6%
	1 - 5 years	13	59.1%
	6 - 10 years	3	13.6%
	11 - 15 years	4	18.2%
	16 - 20 years	2	9.1%
	Over 30 years	0	0.0%
Employer			
	Academia	7	28.0%
	Federal government employee	4	16.0%
	Private sector company	2	8.0%
	State government employee	2	8.0%
	Other	10	40.0%
Job Function			
	Cybersecurity Administrator	0	0.0%
	Cybersecurity Analyst	5	20.0%
	Cybersecurity Architect	2	8.0%
	Cybersecurity Consultant	6	24.0%
	Cybersecurity Engineer	4	16.0%
	Information Assurance Engineer	0	0.0%
	Information Security Analyst	0	0.0%
	Information Security Manager	2	8.0%
	Network Security Engineer	0	0.0%
	Other	6	24.0%

Phase Three – Data Analysis

In phase three the data collected was exported to Microsoft Excel for initial analysis, answers to each survey question were parsed to identify the count of each variable selected by the SMEs. The demographic data was first assessed. Next, the descriptive analysis was prepared for the top six critical cyber visualization variables and the variable for complex data correlations. Table 24 represents the summary for the

calculated mean (M), percentage (%), and standard deviation (SD). Table 31 represents the summary for the calculated mean (M), percentage (%), and standard deviation (SD) for the value.

The level of satisfaction with variable one, workplace satisfaction had an average of 5.20 with a min = 2.00 and max = 7.00. The level of satisfaction with variable two, change in violation patterns had an average of 4.92 with a min = 1.00 and max = 7.00. The level of satisfaction with variable three, audit log modification had an average of 5.12 with a min = 2.00 and max = 7.00. The level of satisfaction with variable four, changes in data access patterns had an average of 5.28 with a min = 1.00 and max = 7.00. The level of satisfaction with variable five, data exfiltration had an average of 5.40 with a min = 2.00 and max = 7.00. The level of satisfaction with variable six, privilege change had an average of 5.60 with a min = 2.00 and max = 7.00. The level of satisfaction with complex cyber data correlations had an average of 4.96 with a min = 1.00 and max = 7.00.

The value of variable one, workplace satisfaction had an average of 5.92 with a min = 2.00 and max = 7.00. The value of variable two, change in violation patterns had an average of 6.12 with a min = 3.00 and max = 7.00. The value of variable three, audit log modification had an average of 6.00 with a min = 3.00 and max = 7.00. The value of variable four, changes in data access patterns had an average of 6.48 with a min = 5.00 and max = 7.00. The value of variable five, data exfiltration had an average of 6.52 with a min = 5.00 and max = 7.00. The value of variable six, privilege change had an average of 6.24 with a min = 4.00 and max = 7.00. The value of complex cyber data correlations had an average of 6.16 with a min = 3.00 and max = 7.00. Table 24 represents the descriptive statistics for the critical cyber visualization variables.

Table 24

Descriptive Statistics for Critical Cyber Visualization Variables (N=25)

Item	***Satisfaction***		***Value***	
	M	***SD***	***M***	***SD***
Variable 1: Workplace Satisfaction	5.20	1.38	5.92	1.22
Variable 2: Change in Violation Patterns	4.92	1.80	6.12	1.13
Variable 3: Audit Log Modification	5.12	1.39	6	1.08
Variable 4: Changes in Data Access Patterns	5.28	1.62	6.48	0.77
Variable 5: Data Exfiltration	5.40	1.58	6.52	0.71
Variable 6: Privilege Change	5.60	1.44	6.24	1.01

Then the data for the level of satisfaction and the value was analyzed in order to determine the respective perceived effectiveness. To address research question five, what is the SMEs' perceived effectiveness (i.e. satisfaction & value/importance) of the QUICK.v™ prototype when mitigating potentially malicious cyber insider threats? After viewing the developed prototype QUICK.v™ SMEs' were presented a 7-point Likert-type rating scale for satisfaction and value to assess for each stated item. Since the statements within survey instrument three were not positively or negatively termed the data was not reverse coded prior to analysis. The satisfaction and value/importance of each item was calculated to then determine the perceived effectiveness using the LeVIS index (Levy, 2006). Based on the LeVIS index the perceived effectiveness was determined using the formula below. Where 49 is used as n to normalize the effectiveness output. This is based on the maximum value and satisfaction scale being seven. Thus,

seven multiplied seven by seven results in the value 49 used to determine the effectiveness in the formula below.

$$\left(\frac{1}{n}\right) * V_0 * S_0 \rightarrow \left(\frac{1}{49}\right) * V_0 * S_0$$

Table 32 represents the perceived effectiveness (LeVIS Index) for all items based on the SMEs responses.

Table 25

LeVIS Index Results for Perceived Effectiveness (N= 25)

Item	Satisfaction	Value	LeVIS Index
Variable 1: Workplace Satisfaction	4.99	5.75	0.59
Variable 2: Change in Violation Patterns	4.42	6.00	0.54
Variable 3: Audit Log Modification	4.91	5.89	0.59
Variable 4: Changes in Data Access Patterns	4.92	6.43	0.65
Variable 5: Data Exfiltration	5.11	6.48	0.68
Variable 6: Privilege Change	5.37	6.15	0.67
Complex Cyber Data Correlations	4.44	6.05	0.55
Type of Variables Presented	4.46	5.58	0.51
Interest in Variables Presented	4.41	5.70	0.51
Organization of Variables Presented	4.32	5.55	0.49
Complexity Based on Variables Presented	4.21	5.44	0.47
Various Variables Were Well Integrated	4.18	5.37	0.46
Relevance of Variables to Insider Threat Detection	4.51	6.31	0.58
Quality of Visualizations	4.10	6.03	0.50
Organization of Visualizations Presented	3.90	5.56	0.45
Consistency of Visualizations Presented	4.41	6.03	0.57
Ability to Quickly Decipher Potential Insider Threats	3.56	6.13	0.45
Confidence Quickly Deciphering Potential Insider Threats	3.88	6.43	0.51
Ability to Make Actionable Decisions Based on Information Depicted	3.83	6.23	0.49
Ease of Use of Information Depicted	4.21	6.21	0.53
Overall, how would you rate your level of satisfaction/importance of QUICK.v™ when identifying potentially malicious cyber insiders?	4.23	5.73	0.49

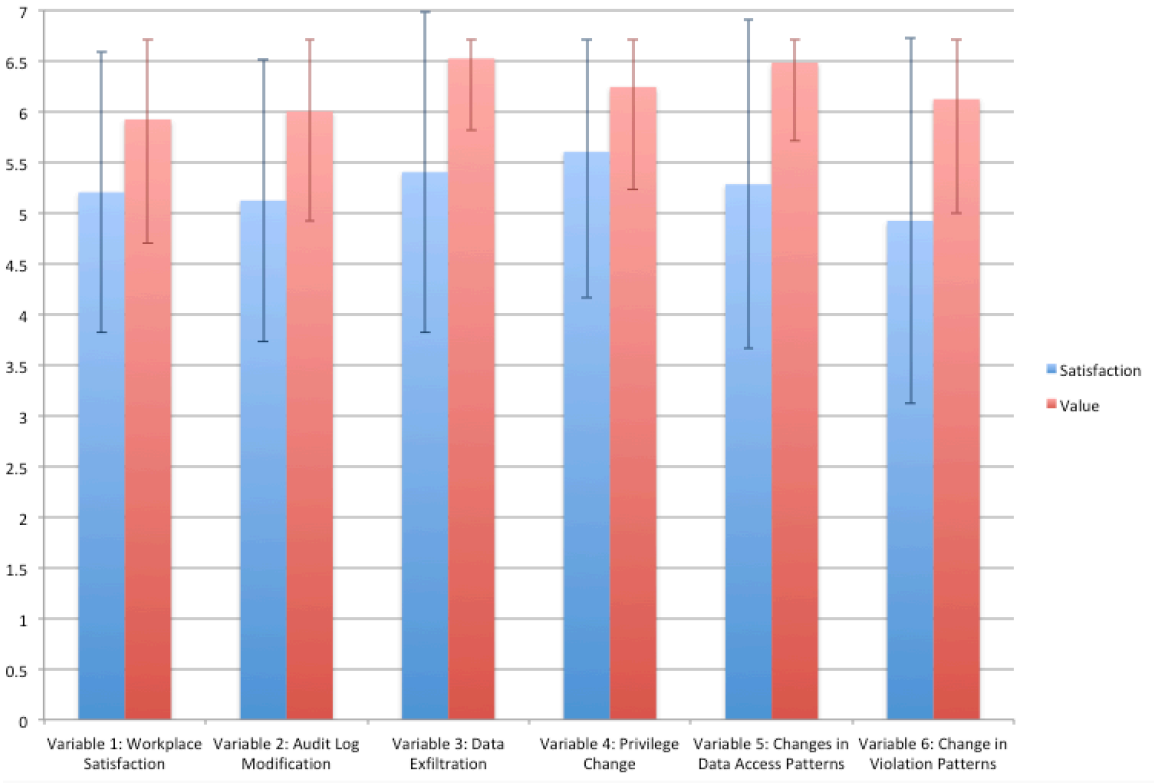


Figure 12: Satisfaction and Value Distribution Summary

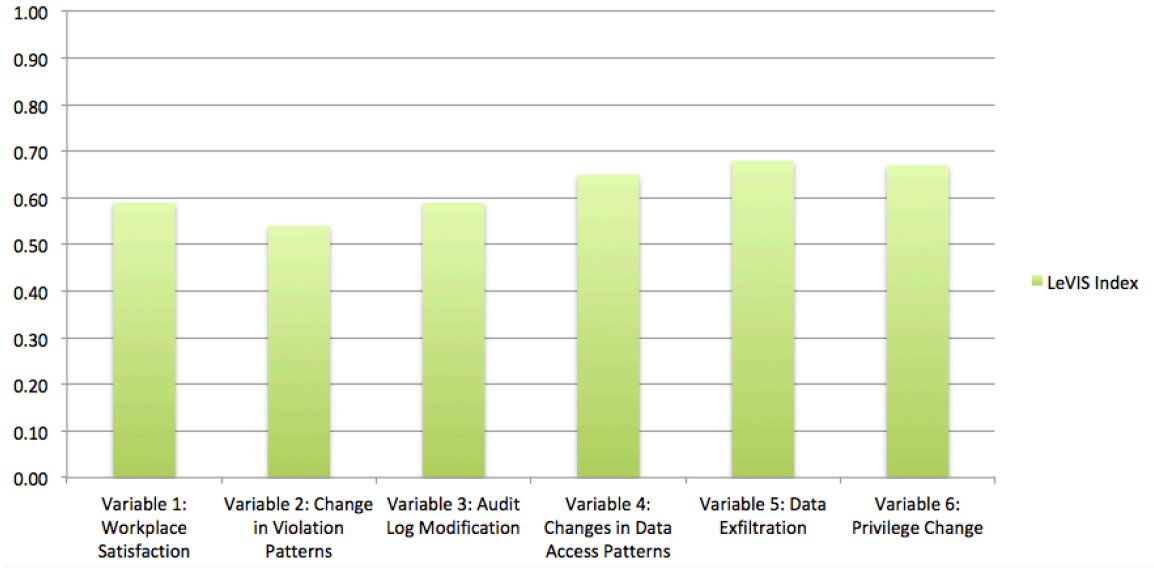


Figure 13: LeVIS Index Summary

Additionally, the modified SUS statements were extracted for analysis. To provide an adjective rating that correlates with the acceptable SUS score of 70 or above. The extracted statements were not framed to alternate between positive and negative statements. Thus, a raw SUS score was calculated based solely on the extracted five SUS statements within survey instrument three. When statements alternate between the positive and negative, care must be taken when scoring the survey (Bangor et al., 2008). The SMEs satisfaction rating for each item was isolated for this analysis. The SUS score was calculated by creating a sum from the items rather than a mean score, this allows for analysis of the same variance as performed by Bangor et al. (2008). Table 26 represents the SUS score for the five items based on all SMEs responses. Table 27 represents the SUS quartile and its corresponding adjective rating based on all SMEs responses.

Table 26

SUS Scores (N=25)

Participant	Level of Satisfaction with "Complexity Based on Variables Presented"	Level of Satisfaction with "Various Variables Were Integrated Well"	Level of Satisfaction with "Consistency (accuracy) of Visualizations Presented"	Level of Satisfaction with "Confidence Quickly Deciphering Potential Insider Threats"	Level of Satisfaction with "Ease of Use of Information Depicted"	Raw SUS Sum Score	Raw SUS Mean Score	Inflated Score (adjusted to a range of 0-100)
p1	4	3	2	5	5	19	3.8	54.3
p2	6	6	7	5	6	30	6	85.7
p3	1	1	1	1	4	8	1.6	22.9
p4	4	6	6	3	4	23	4.6	65.7
p5	6	5	6	6	5	28	5.6	80.0
p6	6	6	6	6	6	30	6	85.7
p7	6	6	6	6	7	31	6.2	88.6
p8	7	7	7	7	7	35	7	100.0
p9	1	1	2	2	2	8	1.6	22.9
p10	4	5	5	4	5	23	4.6	65.7
p11	5	5	7	1	1	19	3.8	54.3
p12	4	5	6	3	3	21	4.2	60.0
p13	2	2	1	1	5	11	2.2	31.4
p14	6	7	5	3	5	26	5.2	74.3
p15	5	3	6	6	6	26	5.2	74.3
p16	5	5	7	6	4	27	5.4	77.2
p17	6	6	6	6	5	29	5.8	82.9
p18	6	6	5	6	6	29	5.8	82.9
p19	4	3	4	4	4	19	3.8	54.3
p20	2	3	3	2	3	13	2.6	37.2
p21	4	4	6	5	5	24	4.8	68.6
p22	7	5	3	7	7	29	5.8	82.9
p23	4	4	5	5	1	19	3.8	54.3
p24	6	6	6	6	5	29	5.8	82.9
p25	6	6	6	6	5	29	5.8	82.9

Table 27

SUS Score by Quartile, Adjective Rating, and Acceptability *(N=25)*

Participant	Inflated Score (adjusted to a range of 0-100)	SUS Quartile	Adjective
p1	54.3	1	Ok
p2	85.7	4	Best Imaginable
p3	22.9	1	Worst Imaginable
p4	65.7	2	Good
p5	80.0	4	Excellent
p6	85.7	4	Best Imaginable
p7	88.6	4	Best Imaginable
p8	100.0	4	Best Imaginable
p9	22.9	1	Worst Imaginable
p10	65.7	2	Good
p11	54.3	1	Good
p12	60.0	1	Good
p13	31.4	1	Poor
p14	74.3	3	Excellent
p15	74.3	3	Excellent
p16	77.2	3	Excellent
p17	82.9	4	Excellent
p18	82.9	4	Excellent
p19	54.3	1	Good
p20	37.2	1	Poor
p21	68.6	2	Good
p22	82.9	4	Excellent
p23	54.3	1	Good
p24	82.9	4	Excellent
p25	82.9	4	Excellent

Phase Three - Comments

No exclusions were identified within the 25 completed surveys. In survey instrument three, SMEs were not asked for comments within the survey. Since the identified list of SMEs consisted of 300 individuals, the surveys were sent to all 300 SMEs to complete on their own. With phase one and two consisting of Delphi survey instruments, by phase three there was an issue of non-response from many SMEs. It is common within Delphi investigations to be unable achieve and maintain an ideal response rate due to the characteristics of multiple iterations, the possible scarcity of

qualified subjects, and the relatively small number of subjects used (Hsu, & Sandford, 2007, p. 1). As a result, only 25 completed surveys were received. From the 25 completed surveys the results indicated each of the identified individual critical cyber visualization variables were effective.

Summary of the Results

In phase one, when SMEs were asked to select the relevant variable that were most important when identifying potentially malicious insider threats, the order in which the variable was selected corresponded to a weight. The weight was used to identify the most critical variable within the analytic variable category. For instance, changes in data access pattern was identified as the second most critical system analytic category variable though, the count equaled twelve. Unauthorized data access methods with the count equal to thirteen, had a significantly lower ranking. Since, the count was equal to seven for variable one which corresponded to changes in data access patterns. The count was equal to three for variable one, unauthorized data access methods. The higher weighting of a selection as variable one resulted in a higher weighted average ranking, consequently increasing the final ranking for changes in data access patterns. Subsequently, in phase one when SME's were asked to select the relevant variable that is most important when trying to identify potentially malicious insider threats within each individual category the findings are as follows.

For the individual categories of system analytic variables, access inconsistent with user class was identified as the most critical variable within this category. In each individual category for facility analytic variables, time of access pattern changes was

identified as the most critical variable within this category. Malware deployment was identified as most critical for the individual category of business capability. Unauthorized or inappropriate associations were identified as the critical visualization variable within the social analytic individual category. For the health individual category, mental instability was identified as a critical visualization variable. Complaints against the user were identified as a critical visualization variable within the human resources individual category. Observed change in means relative to peers was identified as a critical visualization variable for the financial individual category. Unauthorized or inappropriate use of tools was identified as the critical visualization variable for the security individual category. For the criminal analytic variable recent increase in criminal events was identified as the critical visualization variable. These finding vary from the identified rank order of the SMEs' identified critical cyber visualization variables.

When SMEs were given the full list of cyber visualization variables and asked to consider only the top five variables important when identifying potentially malicious insider threats, alternate variables were identified as being most critical. Outside the context of an individual category, SMEs were able to truly focus on what was most critical to them based on their experience. From the ranking of variables a corresponding weight was applied to each variable. The weight was used to identify the top six most critical variables. The identified critical cyber visualization variables were: workplace satisfaction, change in violation patterns, audit log modification, changes in data access patterns, and privilege change.

In phase two, once the SMEs identified and validated the top six critical cyber visualization variables, a comprehensive review of literature was performed to identify

how each variable is visualized. From this review three visualization techniques were selected to be utilized as a visualization technique within phase two. A visual representation of the cyber visualization variable was then mocked up using each visualization technique. SMEs were then presented three images and asked to select the visualization technique most relevant to displaying data related to the critical cyber visualization variable. The results for each critical cyber visualization variable are as follows.

For workplace satisfaction, when presented with the options of a line graph, bar graph, and calendar style view, SMEs identified the line graph as the most valid visualization technique. When presented with an area chart, radar plot, and streamgraph for change in violation pattern, SMEs identified the area chart as the most valid visualization technique. For audit log modification, when presented with a line graph, fisheye distortion, and bar graph, SMEs identified the line graph as the most valid visualization technique. For the critical cyber visualization variable, change in violation pattern when presented with a stacked column graph, a stacked bar graph, and a streamgraph, SMEs identified the stacked column graph as the most valid visualization technique. For data exfiltration, when presented with a line graph, column graph, and fisheye distortion, SMEs identified the line graph as the most valid visualization technique. When presented with a line graph, stacked bars, and stacked columns for the critical cyber visualization variable, privilege change, SMEs identified the line graph as the most valid visualization technique.

To identify visualization techniques are most valid to present complex cyber data correlations, the SMEs were presented with three images showing the relationship

between two or more data variables. The visualization techniques used to present complex cyber data correlations were parallel coordinates, chord diagram, and hierarchical bundling. SMEs identified parallel coordinates as the most valid visualization technique. Since the visualization method identified only varied from the next option by one response, the focus group was sought for additional input. A parallel coordinate was identified as allowing for clear identification of scores. If an SME were in a critical situation and needed to make decisions quickly, a parallel coordinate would allow for faster identification of an issue versus a chord diagram.

In phase three, the perceived effectiveness for the top six critical cyber visualization variables was determined using the Value-Satisfaction Dimension Grid. Table 32 represents the overall effectiveness for all the validated items. Within the Value-Satisfaction Dimension Grids (Figures 14-19) the mean satisfaction scores are on the horizontal axis while the mean value scores are on the vertical axis. The results of the Value-Satisfaction Dimension Grid for workplace satisfaction (Figure 14) indicate that the SMEs perceived effectiveness dimensions are high-value-high-satisfaction. This implies that the critical cyber visualization variable, workplace satisfaction is of high importance to SMEs and the SMEs had high satisfaction with this variable. The results of the Value-Satisfaction Dimension Grid for change in violation patterns (Figure 15) indicate that the SMEs perceived effectiveness dimensions are high-value-high-satisfaction. This implies that the critical cyber visualization variable, change in violation patterns is of high importance to SMEs and the SMEs had high satisfaction with this variable. The results of the Value-Satisfaction Dimension Grid for audit log modifications (Figure 16) indicate that the SMEs perceived effectiveness dimensions are

high-value-moderate-satisfaction. This implies that the critical cyber visualization variable, audit log modifications is of high importance to SMEs and the SMEs had moderate satisfaction with this variable.

The results of the Value-Satisfaction Dimension Grid for change in data access patterns (Figure 17) indicate that the SMEs perceived effectiveness dimensions are high-value-moderate-satisfaction. This implies that the critical cyber visualization variable, change in data access patterns is of high importance to SMEs and the SMEs had moderate satisfaction with this variable. The results of the Value-Satisfaction Dimension Grid for data exfiltration (Figure 18) indicate that the SMEs perceived effectiveness dimensions are high-value-moderate-satisfaction. This implies that the critical cyber visualization variable, data exfiltration is of high importance to SMEs and the SMEs had moderate satisfaction with this variable. The results of the Value-Satisfaction Dimension Grid for privilege change (Figure 19) indicate that the SMEs perceived effectiveness dimensions are high-value-moderate-satisfaction. This implies that the critical cyber visualization variable, privilege change is of high importance to SMEs and the SMEs had moderate satisfaction with this variable.

Table 28

LeVIS Index Results for Perceived Effectiveness Summary

Item	Perceived as Effective?
Variable 1: Workplace Satisfaction	Yes
Variable 2: Change in Violation Patterns	Yes
Variable 3: Audit Log Modification	Yes
Variable 4: Changes in Data Access Patterns	Yes
Variable 5: Data Exfiltration	Yes
Variable 6: Privilege Change	Yes
Complex Cyber Data Correlations	Yes
Type of Variables Presented	Yes
Interest in Variables Presented	Yes
Organization of Variables Presented	No
Complexity Based on Variables Presented	No
Various Variables Were Well Integrated	No
Relevance of Variables to Insider Threat Detection	Yes
Quality of Visualizations	Yes
Organization of Visualizations Presented	No
Consistency of Visualizations Presented	Yes
Ability to Quickly Decipher Potential Insider Threats	No
Confidence quickly Deciphering Potential Insider Threats	Yes
Ability to Make Actionable Decisions Based on Information Depicted	No
Ease of Use of Information Depicted	Yes
Overall, how would you rate your level of satisfaction/value of QUICK.v™ when identifying potentially malicious cyber insiders?	No

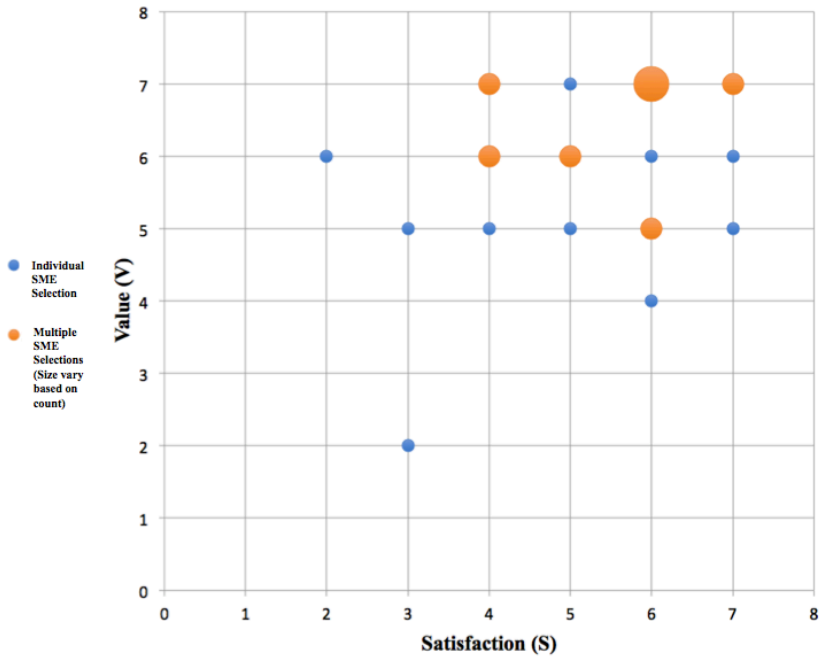


Figure 14: Value-Satisfaction Dimension Grid
Critical Cyber Visualization Variable 1: Workplace Satisfaction

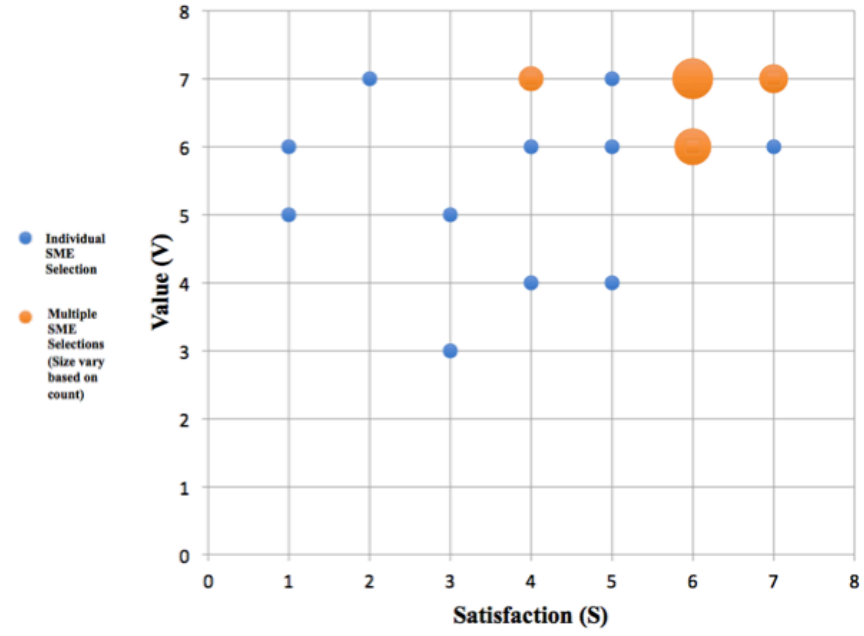


Figure 15: Value-Satisfaction Dimension Grid
Critical Cyber Visualization Variable 2: Change in Violation Patterns

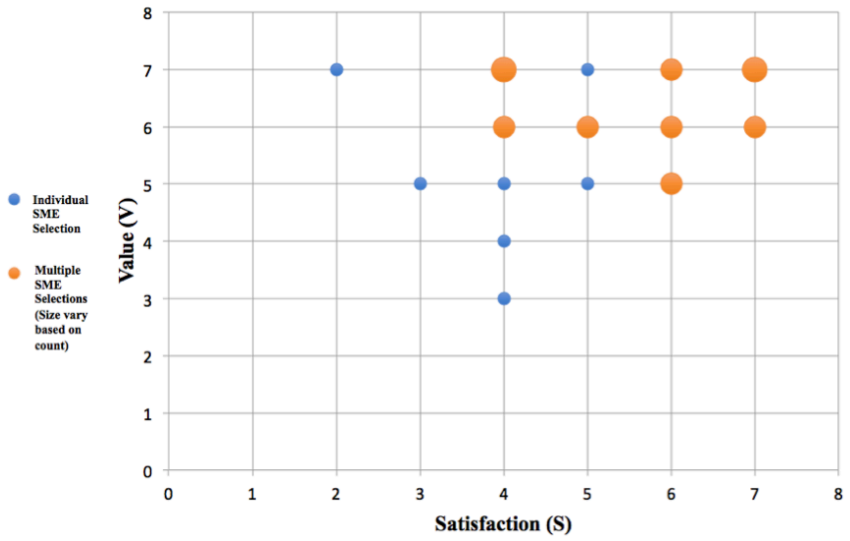


Figure 16: Value-Satisfaction Dimension Grid Critical Cyber Visualization Variable 3: Audit Log Modification

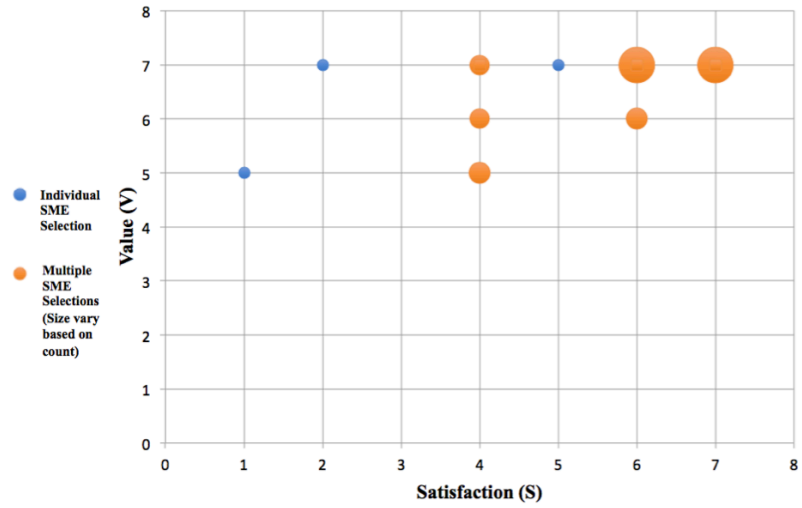


Figure 17: Value-Satisfaction Dimension Grid Critical Cyber Visualization Variable 4: Change in Data Access Patterns

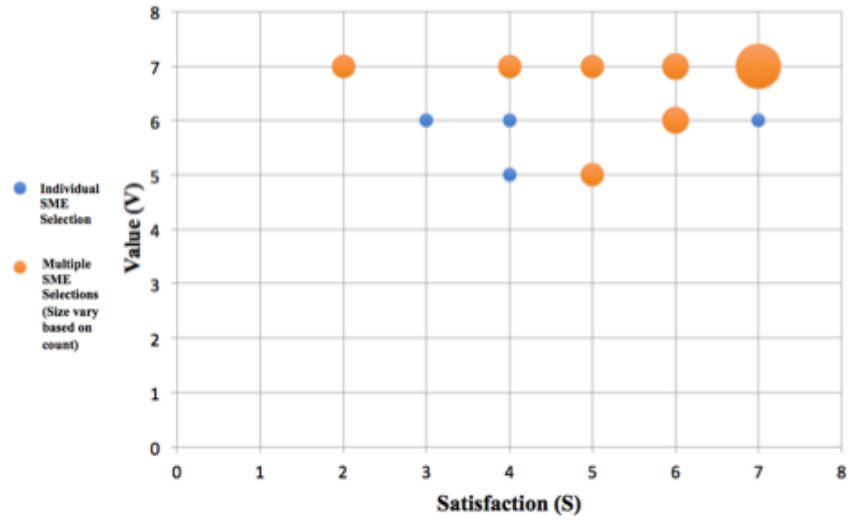


Figure 18: Value-Satisfaction Dimension Grid
Critical Cyber Visualization Variable 5: Data Exfiltration

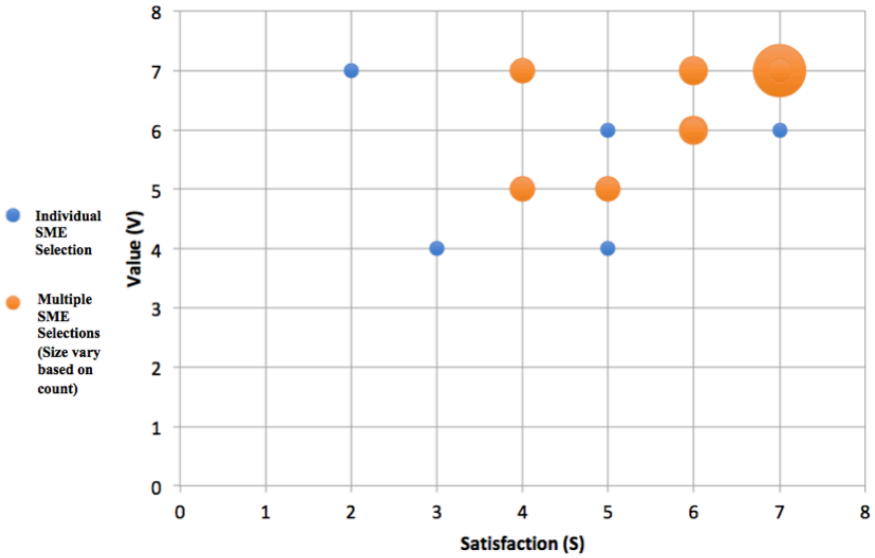


Figure 19: Value-Satisfaction Dimension Grid
Critical Cyber Visualization Variable 6: Privilege Change

Chapter 5

Conclusions, Implications, Recommendations, and Summary

Conclusions

Financial and intellectual property damages continue to rise as a result of insider threats (Cole, 2015; Gorg et al., 2013; Inibhunu et al., 2016; Pfleeger & Stolfo, 2009). This study attempted to address the prevalent challenge within the cybersecurity industry when detecting potentially malicious insider cyber threats and enabling the visualization of threats as they occur. This process was conducted by developing a cyber visualization prototype using SME validated critical cyber visualization variables and techniques. This study achieved the five goals using a three-phased approach. First, using the Dephi method SMEs identified and ranked the critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats. Next, using the Delphi method SMEs identified visualization techniques that are most valid to present complex cyber data correlations and the top six critical cyber visualization variables. Finally, SMEs identified the perceived effectiveness of the developed prototype QUICK.v™.

Discussion

Overall, the results of the study designated the top six critical cyber visualization variables: workplace satisfaction, change in violation patterns, audit log modification, changes in data access patterns, data exfiltration, and privilege change. These results

suggest that cybersecurity analysts should initially focus on anomalies within these areas when using applications to detect potentially malicious insider cyber threats. The results also indicated most valid visualization technique to present complex cyber data correlations are parallel coordinates. The most valid visualization technique to present the top six critical cyber visualization variables are: line graph, area chart, line graph, stacked column graph, line graph, and line graph (denoted in the order that each variable was previously listed). The results also identified parallel coordinates as the most valid to present complex cyber data correlations relevant to the pre-designated critical cyber visualization variables. This suggests that cybersecurity analysts should be presented simplified visualizations using these visualization techniques when presenting the critical cyber visualization variables.

Overall, QUICK.v™ was not implied to be effective based on the SMEs' overall perceived effectiveness rating (i.e. satisfaction & value/importance) when mitigating potentially malicious cyber insider threats. However, each of the identified individual critical cyber visualization variables was found to be effective. Perceived effectiveness was also implied for the following items: complex cyber data correlations, the types of variables presented, interest in the variables presented, relevance of the variables to insider threat detection, quality of visualizations, consistency of visualizations presented, confidence quickly deciphering potential insider threats, and ease of use of information depicted. Perceived effectiveness was not implied for the following items: organization of variables presented, complexity based on variables presented, various variables were well integrated, organization of visualizations presented, ability to quickly decipher potential

insider threats, ability to make actionable decisions based on information depicted, and the overall value and satisfaction of QUICK.v™.

From the five items analyzed to determine the SUS scores, thirteen of the 25 participants had an SUS score above 70, which is deemed as acceptable. The sample average SUS score was 66.9%. Thus the overall perceived usability or user satisfaction of QUICK.v™ based on the five modified SUS statements fell within quartile two which is deemed as satisfied based on the SUS Score by quartile, adjective rating, and acceptability (Bangor et al., 2008).

Implications

There are implications of this study in relation to the existing body of knowledge in IS and InfoSec. This study developed a novel and effective detection method for the identification of anomalous activities when mitigating malicious insider cyber threats. Many cybersecurity tools presenting visualizations are rarely evaluated for effectiveness nor do they account for the needs of the user (Sethi et al., 2016). This study identified SME validated cybersecurity vital signs, their corresponding visualization technique, and validated the effectiveness. Since financial and intellectual property loses continue to rise due to insider threats, it is important to ensure cybersecurity analysts are enabled to mitigate potentially malicious cyber insider threats.

In this study, the SME data collection occurred over a span of fifteen weeks. This time period allowed for SME participation and follow-up. Using an expert panel required continual follow-up, which resulted in delays. Though follow-ups were found to be a method for reducing non-response within the Delphi process, a drawback to Delphi

process is that the questionnaire method may slow data collection (Chang et al., 2018).

This study provides companies with cybersecurity vital signs that are perceived as effective when identifying potentially malicious cyber insiders. These cybersecurity vital signs could assist with the identification and mitigation of malicious insiders cyber threats.

Recommendations and Future Research

Generally, when asked to determine criticality while restricted within one category, SMEs identified with variables that they knew they could most likely measure or determine based on data. Thus, workplace satisfaction, a variable identified as most critical fell to a lowered ranking within the individual category as most SMEs acknowledged this as data they would not be able to obtain. The SMEs then selected a choice that logically seemed more attainable to them like unauthorized or inappropriate associations. Since this data may be possibly obtained from an employees social media connections. Workplace satisfaction though critical was ranked lower since the data seemed unattainable. In giving the SMEs all 45 variables to parse and select only five, they subjectively chose what seemed most pertinent without over analysis. Future studies would increase the validity of the identified critical cyber visualization variables.

When asked to identify valid visualization techniques though SMEs were drawn to the more unique visualization options, they opted to select the more simple visualization techniques. These visualization techniques were identified with being more familiar to the SMEs. It can be denoted that the identification with more simplified visualization techniques, in the form of lines and charts, reflects similarities to the type of

visualizations standardized on EKGs. Like medical professionals cybersecurity professionals may prefer simplified visualizations. It may be assumed that simplified visualizations may reduce cognitive load during times of crisis. However, this assumption may be further investigated within future research.

Summary

This study addressed the prevalent challenge faced within the cybersecurity industry when detecting potentially malicious insider cyber threats, to enable visualization of those threats as they occur (Gorg et al., 2013; Inibhunu et al., 2016; Pfleeger & Stolfo, 2009; Patcha & Park, 2007). Insider threat, threatens personal data, national security, as well as economic prosperity (Pfleeger et al., 2010). In cybersecurity mitigating adverse incidents require surveillance to identify anomaly metrics and attack patterns (Agrafiotis et al., 2015). Delays in identifying a potentially malicious cyber insider may result in substantial losses, resulting in the deterioration of an organization (Randazzo et al., 2005).

Detecting malicious cybersecurity insiders is a complex task since their malicious actions take place alongside normal behavior (Azaria, Richardson, Kraus, & Subrahmanian, 2014). Detecting misuse by malicious cybersecurity insiders involves examining an individual's use of information and resources to determine whether the use is appropriate (Caputo, Maloof, & Stephens, 2009). Identifying malicious behaviors amidst appropriate activities pose an issue. Cybersecurity analysts must identify outliers or anomalies within a users activities utilizing information visualization (Kang & Gorg, 2011). Therefore, identifying anomalies and taking corrective action pose a challenge

(Azaria et al., 2014).

The main goal of this research was to validate empirically a dashboard visualization prototype for cross team collaboration and proactive responses when reacting to malicious cybersecurity insiders. Building on the works of Albanese, Pugliese, and Subrahmanian (2013), Boukri and Chaoui (2015), Greitzer and Hohimer (2011), Legg et al. (2015), as well as Legg, Moffat, Nurse, and Happa (2013). Albanese et al. (2013) developed a graphical index that would aid in providing evidence of occurrences of an activity, and identify if a problem matches a sequence of observations. Though the threat posed by malicious cybersecurity insiders is very real, there is a lack of actual analysis of activity logs due to the sheer volume of activities being conducted daily (Legg et al., 2015).

This study had five specific goals. The first research goal was to identify the critical cyber visualization variables. The second research goal was to identify the rank order of the critical cyber visualization variables that the developed prototype should include, which may aid in identifying potentially malicious cyber insiders. The third research goal was to identify the most valid presentation of complex data correlations using the identified critical visualization variables over multiple visualization techniques. The fourth research goal was to apply SMEs' identified critical visualization variables, in rank order, and techniques to develop QUICK.v™. The fifth research goal was to conduct an experimental study to assess the perceived effectiveness using self-reported value and satisfaction of the QUICK.v™ prototype when mitigating malicious cyber insiders.

In phase one, an exploratory study was conducted. Cybersecurity SMEs were solicited from the researches LinkedIn professional contacts to answer the following research questions:

RQ1: What are SMEs' identified critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats?

RQ2: What is the rank order of the SMEs' identified critical cyber visualization variables that should be displayed when developing a cyber insider threat dashboard visualization prototype to detect potentially malicious cyber insider activities?

The Delphi method was used in order to obtain consensus among SMEs on the identified critical cyber visualization variables and their corresponding rank order. The result of the survey was the identified top six critical cyber visualization variables.

Once the SMEs identified and validated the top six critical cyber visualization variables, a comprehensive review of literature was performed to identify how each variable is visually represented. In phase two, data collection was performed with cybersecurity and visualization SMEs. The SMEs consisted of cybersecurity and visualization professionals. These individuals were sourced from the researchers LinkedIn network. In phase two, SMEs were then presented three images and asked to select the visualization technique most relevant to displaying data related to the critical cyber visualization variable, in order to answer the following research questions:

RQ3: What SMEs' identified visualization techniques are most valid to present complex cyber data *correlations* relevant to the pre-designated critical

cyber visualization variables that are applied within the developed cyber visualization prototype QUICK.v™?

RQ4: What SMEs' identified visualization techniques are most valid to present *top six critical cyber visualization variables* to detect potentially malicious cyber insider activities that are applied within the developed cyber visualization prototype QUICK.v™?

In phase three, qualitative and quantitative data was collected from cybersecurity and visualization professionals. SMEs were asked to view the developed prototype and identify level of satisfaction and importance in order to answer the following research question:

RQ5: What is the SMEs' perceived effectiveness (i.e. satisfaction & value/importance) of the QUICK.v™ prototype when mitigating potentially malicious cyber insider threats?

This study made several contributions to Information Systems and Information Security body of knowledge by developing a novel and effective detection method for the identification of anomalous activities when mitigating malicious insider cyber threats. The study provided empirical evidence regarding the magnitude of endless alerts, increasing the time required for decision-making when identifying potentially malicious insiders cyber threats. Given the significant financial and intellectual property damage posed to organizations, the results of this study provides organizations with empirical evidence of how to visualize cybersecurity vital signs pertinent to the identification of malicious cyber insiders. Lack of identification and mitigation of potentially malicious cyber insider threats could result in substantial financial losses for an organization, or

government entity.

In conclusion, organizations can use the identified cybersecurity vital signs and the validated visualization techniques to aid in the identification of malicious insiders. QUICK.v™ addressed the challenge of detecting cyber insider threats in an unconventional way by enhancing the presentation of complex malicious insider cyber threat correlations. In addition, QUICK.v™ can be used as a guide for alleviating the issues faced when using visualizations to identify potentially malicious insiders cyber threats.

Appendices

Appendix A

Qualitative Survey Instrument 1(TEMPLATE): Instrument for SMEs Identification of Cyber Visualization Variables

Please read the following instructions and definitions before completing this survey

The purpose of this survey is to identify cyber visualization variables (analytic variable or outputs that may indicate an insider threat and prompts for further analysis) that a cyber specific information visualization should include that may aid in identifying potentially malicious insiders cyber threats.

Please respond to all questions as honestly and accurately as possible. By completing this survey you agree and understand that your responses are voluntary. Measures were taken to ensure that responses are anonymous and cannot be traced to any individual. You may exit this survey at any time. In the event that you chose to exit this survey, your responses will not be recorded. By participating in this survey you certify that you are over the age of 18 years old.

Cyberspace or 'Cyber' – Independent network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries (White House, 2009).

Cybersecurity – Prevention of damage to, unauthorized use of, exploitation of, and, if needed, the restoration of electronic information and communications systems to ensure confidentiality, integrity, and availability (Axelrod, 2006).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Insider Threat - Individuals with legitimate access whose behaviors put a firm's data, intellectual property, systems, organizations, and businesses at risk of being attacked (Pfleeger, Predd, Hunker, & Bulford, 2010).

Intrusion Detection System - Hardware or software that gathers and analyzes information from various areas within a computer or a network to identify possible security breaches, which include both intrusions and misuse (NIST, 2013).

Malicious Insider - Is an insider who has malicious intent that acts against the best interests of the organization (Santos et al., 2012).

Section A

Initial Cyber Visualization Variables

From the cyber visualization variables presented below, place an x next to the variables that a most important to you when trying to identify potentially malicious insider threats.

(*Example: Based on the sample list below – 'System authentication and authorization failures')*

Category	Analytic Variables	
Activity-Based Analytics		**Select (x)**
System	Authentication and Authorization Failure	
	Changes in Data Access Patterns	
	Access Inconsistent With User Class	
	Changes in Network Patterns	
	Network Patterns Inconsistent with User Class	
	Data Exfiltration	
	Unauthorized Data Access Methods	
	Privilege Change	
	Erroneous Defensive Posture Changes	
	Improper Command Usage	
	Knowledge Access	
	Audit Log Modification	
Facility	Time of Access Pattern Changes	
	Locality of Access Pattern Changes	
Business Capabilities	Failure Correlation	
	Malware Deployment	
	Deletion or Modification of Data or Infrastructure	
	Analysis of Competitor	
	Analysis of Public Media	
	Attribution of Disclosure	
	Retrieval	

Content-Based Analytics		Select (x)
Social	Disregard	
	Personal Inflexibility	
	Unusual Contacts	
	Unusual Business Travel	
	Unusual Personal Travel	
	Unauthorized or Inappropriate Associations	
	Withdrawal	
	Workplace Events	
	Workplace Satisfaction	
Health	Mental instability	
	Impulse Control	
Human Resources	Major Life Event	
	Complaints Against the User	
	Negative Reviews	

Inferential Analytics		Select (x)
Financial	Observed Temporal Change in Means	
	Observed Change in Means Relative to Peers	
	Financial Reporting	
Security	Change in Violation Patterns	
	Duration and Regularity of Security Events	
	Unauthorized or Inappropriate Use of Tools	
Criminal	Restraining Orders	
	Wage Garnishments, etc.	
	Violence Outside Workplace	
	Recent Increase in Criminal Events	

1. *List any additional* cyber visualization variables *you think will support the identification of potentially malicious insiders cyber threats, that are not listed above.*
   (*Example: Based on the sample list above – 'System authentication and authorization failures')*

   _____

   _____

   _____

   _____

   _____

2. *From all the variables identified above, please think of issues related to insider threat detection and list only the top 5 variables important to you when identifying potentially malicious insider threats.*
    1. _____
    2. _____
    3. _____
    4. _____
    5. _____

3. *From the 5 variables identified in Question 2, please think of issues related to insider threat detection and rank the variables from most important to least important (1 – least important and 5- most important).*
    1. _____
    2. _____
    3. _____
    4. _____
    5. _____

4. *What are the positives of using the 5 variables identified in Question 2 when identifying potentially malicious insiders cyber threats?*

    _____
    _____
    _____
    _____
    _____
    _____
    _____

5. *Are there any negatives associated with using the 5 variables identified in Question 2 when identifying potentially malicious insiders cyber threats?*

    _____
    _____

Section B

Demographic Information

What is your job function?

(A) Administrative staff
(B) Cybersecurity/IT staff
(C) Engineer
(D) Manager
(E) Operations
(F) Professional staff
(G) Security operator
(H) Technical staff
(I) Other _____

How long have you been with your current organization?

(A) Under 1 year
(B) 1 – 5 years
(C) 6 – 10 years
(D) 11 – 15 years
(E) 16 – 20 years
(F) 21 – 25 years
(G) 26 – 30 years
(H) Over 30 years

Which describes your current employer?

(A) Academia
(B) Federal government employee
(C) Private sector company
(D) State government employee
(E) Other _____

What is your highest level of education?

(A) High school diploma
(B) 2-year college (Associates degree)
(C) 4-year college (Bachelors degree)
(D) Graduate degree
(E) Doctorate
(F) Other

Do you currently hold any cybersecurity certifications, if so how many do you possess?

(A) 0
(B) 1
(C) 2
(D) 3
(E) 4
(F) 5 or more

What is your age?

(A) Under 20
(B) 20-29
(C) 30-39
(D) 40-49
(E) 50-59
(F) Over 60

What is your gender?

(A) Female
(B) Male

Appendix B

Qualitative Survey Instrument 2 (TEMPLATE): Instrument for SME Identification of Visualization Technique for Cyber Variables

Please read the following definitions before completing this survey

This Appendix is only a template since the top six critical cyber visualization variables will not be finalized until the completion of Appendix A. Thus, allowing for the identification of the most valid visualization techniques to present the top six critical cyber visualization variables, therefore, this Appendix is only for illustration purposes.

The purpose of this survey is to identify the visualization techniques best suited to present the cyber visualization variables previously identified by SMEs that may aid in identifying potentially malicious insiders cyber threats.

Cyberspace or 'Cyber' – Independent network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries (White House, 2009).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Insider Threat - Individuals with legitimate access whose behaviors put a firm's data, intellectual property, systems, organizations, and businesses at risk of being attacked (Pfleeger, Predd, Hunker, & Bulford, 2010).

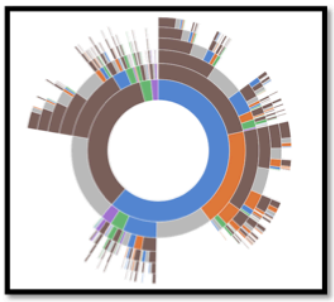
Intrusion Detection System - Hardware or software that gathers and analyzes information from various areas within a computer or a network to identify possible security breaches, which include both intrusions and misuse (NIST, 2013).

Malicious Insider - Is an insider who has malicious intent that acts against the best interests of the organization (Santos et al., 2012).

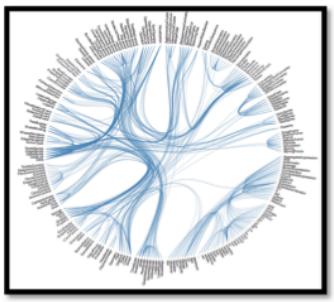
Section A

1. *Please review the visualization technique used to present each of the variables depicted below and select your preferred presentation technique for each variable.*

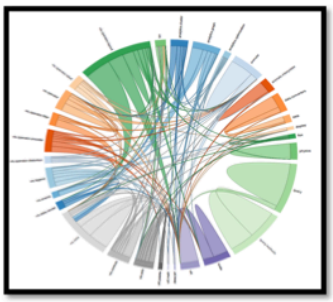
Complex Data Correlations



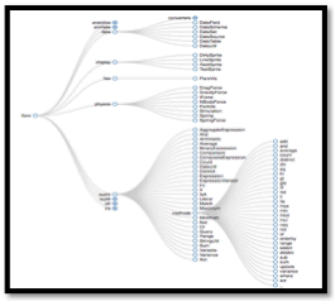
Visualization Technique 1 ☐



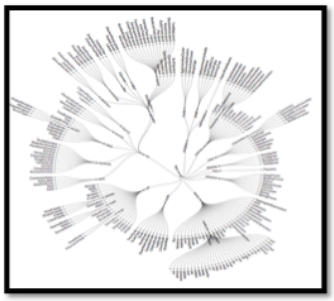
Visualization Technique 2 ☐



Visualization Technique 3 ☐



Visualization Technique 4 ☐

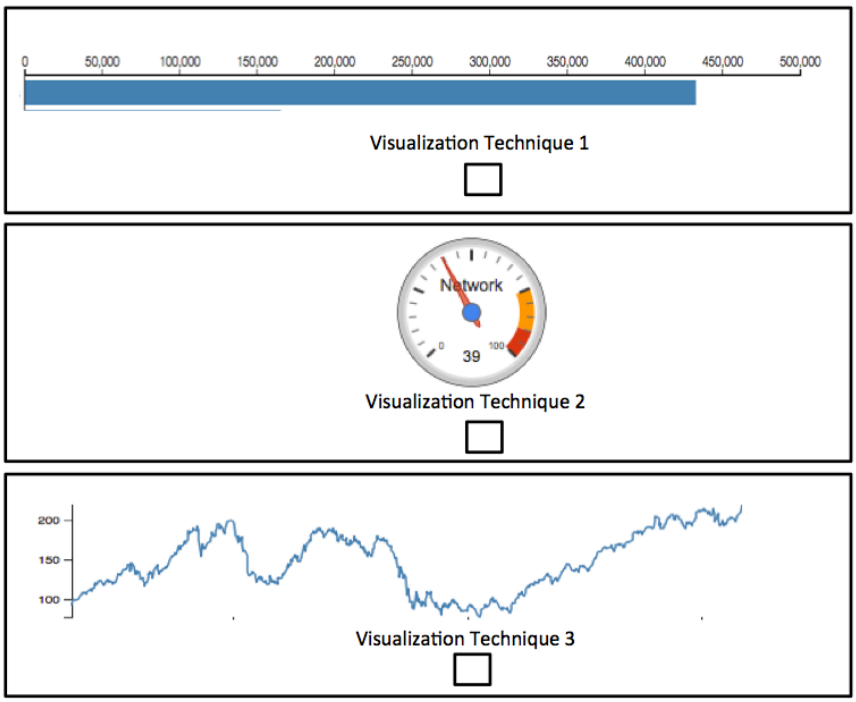


Visualization Technique 5 ☐

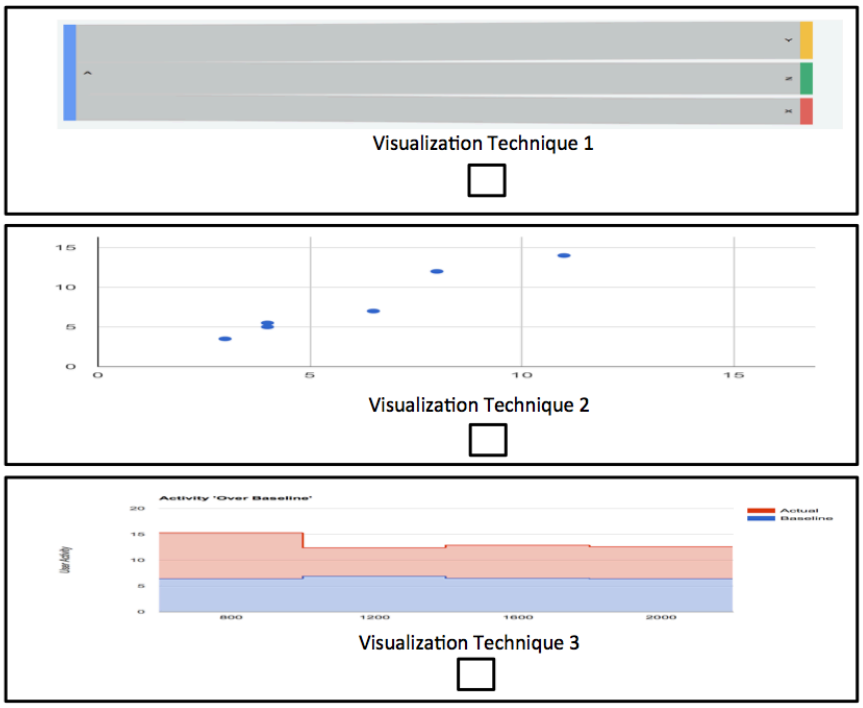
Visualization Technique 6 ☐

Top Five Variables

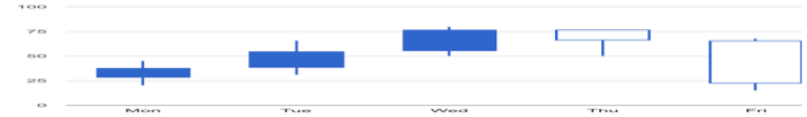
Variable 1:



Variable 2:



Variable 3:



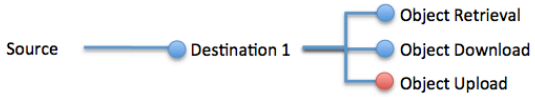
Visualization Technique 1

☐

	Name	% Above Baseline	High Risk Employee
1	Mike	5%	✓

Visualization Technique 2

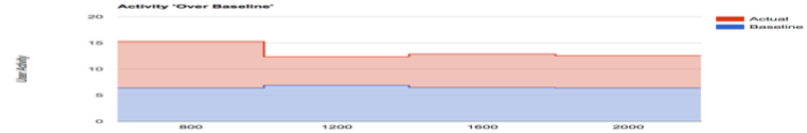
☐



Visualization Technique 3

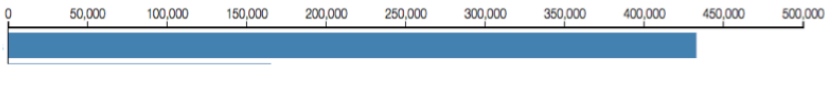
☐

Variable 4:



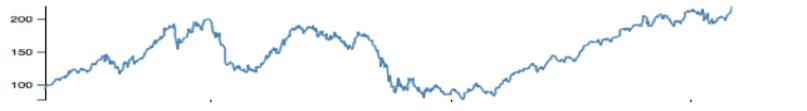
Visualization Technique 1

☐



Visualization Technique 2

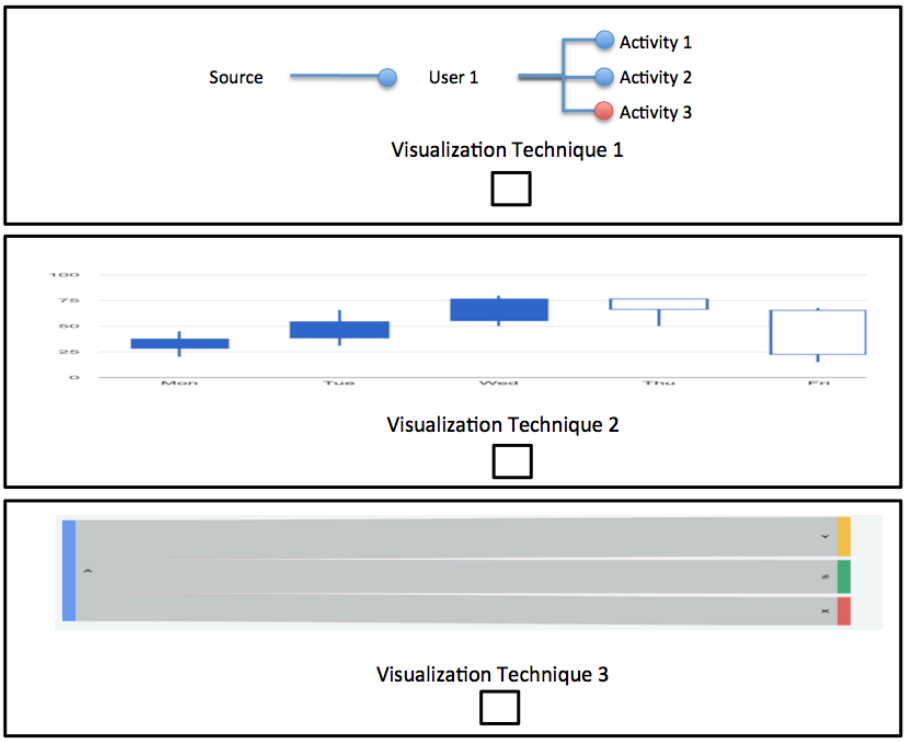
☐



Visualization Technique 3

☐

Variable 5:


Source
User 1
Activity 1
Activity 2
Activity 3
Visualization Technique 1
100
75
50
25
0
Mon
Tue
Wed
Thu
Fri
Visualization Technique 2
Visualization Technique 3

Section B

Demographic Information

What is your job function?

(A) Administrative staff
(B) Cybersecurity/IT staff
(C) Engineer
(D) Manager
(E) Operations
(F) Professional staff
(G) Security operator
(H) Technical staff
(I) Other _____

How long have you been with your current organization?

(A) Under 1 year
(B) 1 – 5 years
(C) 6 – 10 years
(D) 11 – 15 years
(E) 16 – 20 years
(F) 21 – 25 years
(G) 26 – 30 years
(H) Over 30 years

Which describes your current employer?

(A) Academia
(B) Federal government employee
(C) Private sector company
(D) State government employee
(E) Other _____

What is your highest level of education?

(A) High school diploma
(B) 2-year college (Associates degree)
(C) 4-year college (Bachelors degree)
(D) Graduate degree
(E) Doctorate
(F) Other

Do you currently hold any cybersecurity certifications, if so how many do you possess?

(A) 0
(B) 1
(C) 2
(D) 3
(E) 4
(F) 5 or more

What is your age?

(A) Under 20
(B) 20-29
(C) 30-39
(D) 40-49
(E) 50-59
(F) Over 60

What is your gender?

(A) Female
(B) Male

Appendix C

Quantitative Survey Instrument 3 (TEMPLATE): Instrument for Cybersecurity Analysts' Perceived Effectiveness of the Prototype

Please read the following definitions before completing this survey

The purpose of this survey is to determine the perceived effectiveness of the cyber visualization variables presented and the visualization techniques used was identified. The instrument depicted below is a draft for the final version that was developed based on SME input. Keep the definitions below in mind as you complete the survey.

Cyberspace or 'Cyber' – Independent network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries (White House, 2009).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Insider Threat - Individuals with legitimate access whose behaviors put a firm's data, intellectual property, systems, organizations, and businesses at risk of being attacked (Pfleeger, Predd, Hunker, & Bulford, 2010).

Intrusion Detection System - Hardware or software that gathers and analyzes information from various areas within a computer or a network to identify possible security breaches, which include both intrusions and misuse (NIST, 2013).

Malicious Insider - Is an insider who has malicious intent that acts against the best interests of the organization (Santos et al., 2012).

Section A

Instrument for Cybersecurity Analysts Effectiveness

The following lists are related to the cybersecurity variables and visualizations used within the developed prototype QUICK.v™. Please read each item and rate your level of satisfaction and level of importance you attribute to each item when detecting potentially malicious cybersecurity insider threats. Rate the level of satisfaction for each item from: 'Extremely Unsatisfied' to Extremely Satisfied'. Also, rate each items level of importance from: 'Not Important' to Extremely Important'.

A. The following items relate to the cybersecurity variables presented:

Items	Level of Satisfaction							Level of Importance						
	Extremely Unsatisfied	Very Unsatisfied	Unsatisfied	Neutral	Satisfied	Very Satisfied	Extremely Satisfied	Not Important	Not so Important	Slightly Important	Neutral	Important	Very Important	Extremely Important
A1. Variable 1	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A2. Variable 2	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A3. Variable 3	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A4. Variable 4	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A5. Variable 5	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A6. Type of variables presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A7. Interest in variables presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A8. Organization of variables presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A9. Complexity based on variables presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A10. Various variables were well integrated	○	○	○	○	○	○	○	○	○	○	○	○	○	○
A11. Relevance of variables to insider threat detection	○	○	○	○	○	○	○	○	○	○	○	○	○	○

B. The following items relate to the visualization methods used to present the cybersecurity variables:

Items	Level of Satisfaction							Level of Importance						
	Extremely Unsatisfied	Very Unsatisfied	Unsatisfied	Neutral	Satisfied	Very Satisfied	Extremely Satisfied	Not Important	Not so Important	Slightly Important	Neutral	Important	Very Important	Extremely Important
B1. Quality of visualizations	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B2. Organization of visualizations presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B3. Consistency of visualizations presented	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B4. Ability to quickly decipher potential insider threats	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B5. Confidence quickly deciphering potential insider threats	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B6. Ability to make actionable decisions based on information depicted	○	○	○	○	○	○	○	○	○	○	○	○	○	○
B7. Ease of use of information depicted	○	○	○	○	○	○	○	○	○	○	○	○	○	○

C. Global Questions:

C1. Overall, how would you rate your level of satisfaction with QUICK.v™ when identifying potentially malicious cyber insiders?	○	○	○	○	○	○	○							
C2. Overall, how important would QUICK.v™ be to you when identifying potentially malicious cyber insiders?								○	○	○	○	○	○	○

Section B

Demographic Information

What is your job function?

(A) Administrative staff
(B) Cybersecurity/IT staff
(C) Engineer
(D) Manager
(E) Operations
(F) Professional staff
(G) Security operator
(H) Technical staff
(I) Other _____

How long have you been with your current organization?

(A) Under 1 year
(B) 1 – 5 years
(C) 6 – 10 years
(D) 11 – 15 years
(E) 16 – 20 years
(F) 21 – 25 years
(G) 26 – 30 years
(H) Over 30 years

Which describes your current employer?

(A) Academia
(B) Federal government employee
(C) Private sector company
(D) State government employee
(E) Other _____

What is your highest level of education?

(A) High school diploma
(B) 2-year college (Associates degree)
(C) 4-year college (Bachelors degree)
(D) Graduate degree
(E) Doctorate
(F) Other

Do you currently hold any cybersecurity certifications, if so how many do you possess?

(A) 0
(B) 1
(C) 2
(D) 3
(E) 4
(F) 5 or more

What is your age?

(A) Under 20
(B) 20-29
(C) 30-39
(D) 40-49
(E) 50-59
(F) Over 60

What is your gender?

(A) Female
(B) Male

Appendix D

Qualitative Survey Instrument 1 (FINAL): Email to SMEs

Cyber Insider Threat 10-15 Minute Survey

Cyber insider threat is one of the most difficult risks to mitigate in organizations. However, effective validated visualizations for cyber analysts to better decipher and react to detected anomalies has yet to be reported in literature or in industry. The main goal of this research study is to develop and validate, using Subject Matter Experts (SME), an executive insider-threat dashboard visualization prototype. The goal of the developed prototype is to alleviate the issues faced when using visualizations to identify potentially malicious cyber insiders.

As an expert in the field, your assistance is needed to identify critical cyber visualization variables that should be displayed when using applications to detect potentially malicious insider cyber threats. The initial items provided in this survey were validated in prior research and peer reviewed literature. However, this is the first time that they will be validated to identify critical cyber visualization variables. The feedback that you provide will be used for this research study in aggregated form. This survey is simple and will take only 10-15 minutes. As an expert in the field, you agree to keep all information regarding this research confidential and not to disclose any details related to this survey or the material contained within it.

I am a Ph.D. candidate in Information Systems at the College of Engineering and Computing, Nova Southeastern University, working under the supervision of Professor Yair Levy (levy@nova.edu), and a member of his Levy CyLab (http://CyLab.nova.edu/). If you have any questions regarding this research, please do not hesitate to contact me or Dr. Levy via email.

Please send this invitation to participate in the survey to friends and colleagues with cybersecurity expertise. Thank you in advance for your time and consideration. I appreciate your assistance and contribution to this research study.

Regards,
Karla A. Clarke, Ph.D. Candidate
Kc1127@mynsu.nova.edu

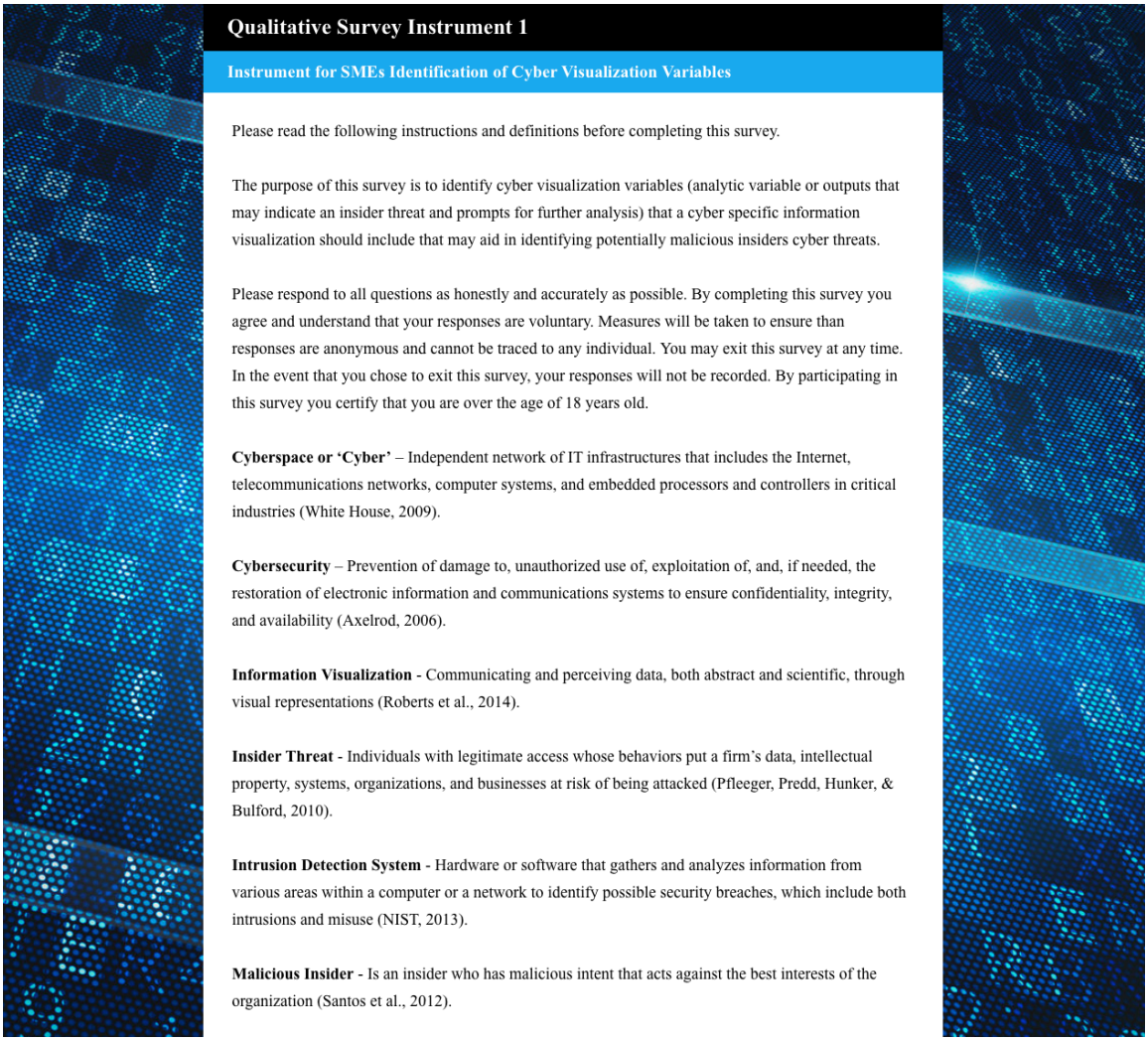
Begin Survey

Please do not forward this email as its survey link is unique to you.
Unsubscribe from this list

Powered by SurveyMonkey

Appendix E

Qualitative Survey Instrument 1 (FINAL): Instrument for SMEs Identification of Cyber Visualization Variables



Qualitative Survey Instrument 1

Instrument for SMEs Identification of Cyber Visualization Variables

Please read the following instructions and definitions before completing this survey.

The purpose of this survey is to identify cyber visualization variables (analytic variable or outputs that may indicate an insider threat and prompts for further analysis) that a cyber specific information visualization should include that may aid in identifying potentially malicious insiders cyber threats.

Please respond to all questions as honestly and accurately as possible. By completing this survey you agree and understand that your responses are voluntary. Measures will be taken to ensure than responses are anonymous and cannot be traced to any individual. You may exit this survey at any time. In the event that you chose to exit this survey, your responses will not be recorded. By participating in this survey you certify that you are over the age of 18 years old.

Cyberspace or 'Cyber' – Independent network of IT infrastructures that includes the Internet, telecommunications networks, computer systems, and embedded processors and controllers in critical industries (White House, 2009).

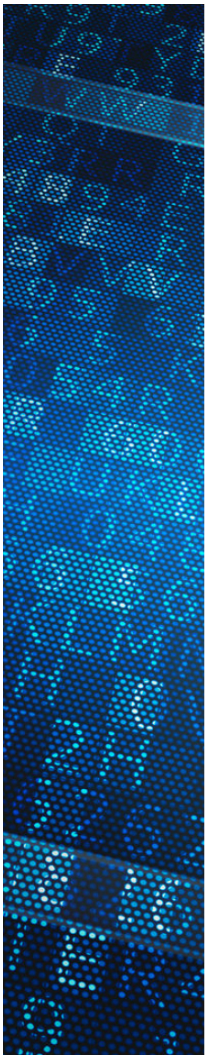
Cybersecurity – Prevention of damage to, unauthorized use of, exploitation of, and, if needed, the restoration of electronic information and communications systems to ensure confidentiality, integrity, and availability (Axelrod, 2006).

Information Visualization - Communicating and perceiving data, both abstract and scientific, through visual representations (Roberts et al., 2014).

Insider Threat - Individuals with legitimate access whose behaviors put a firm's data, intellectual property, systems, organizations, and businesses at risk of being attacked (Pfleeger, Predd, Hunker, & Bulford, 2010).

Intrusion Detection System - Hardware or software that gathers and analyzes information from various areas within a computer or a network to identify possible security breaches, which include both intrusions and misuse (NIST, 2013).

Malicious Insider - Is an insider who has malicious intent that acts against the best interests of the organization (Santos et al., 2012).



1. **System Activity-Based Analytics**

From the variables presented in the dropdown, select the relevant system variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
System		

2. **Facility Activity-Based Analytics**

From the variables presented in the dropdown, select the relevant facility variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Facility		

3. **Business Capabilities Activity-Based Analytics**

From the variables presented in the dropdown, select the relevant business capability variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Business Capabilities		

4. **Social Content-Based Analytics**

From the variables presented in the dropdown, select the relevant social variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Social		

5. **Health Content-Based Analytics**

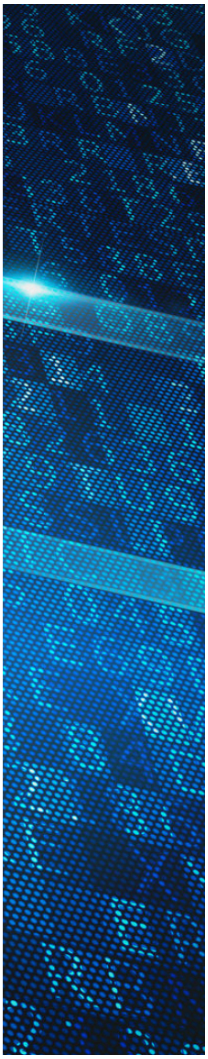
From the variables presented in the dropdown, select the relevant health variable(s) that are most important to you when trying to identify potentially malicious insider threats.

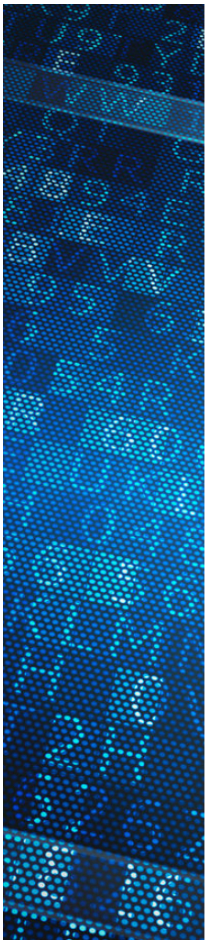
	Analytic Variable 1	Analytic Variable 2
Health		

6. **Human Resources Content-Based Analytics**

From the variables presented in the dropdown, select the relevant human resources variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Human Resources		





7. **Financial Inferential Analytics**

From the variables presented in the dropdown, select the relevant financial variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Financial		

8. **Security Inferential Analytics**

From the variables presented in the dropdown, select the relevant security variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Security		

9. **Criminal Inferential Analytics**

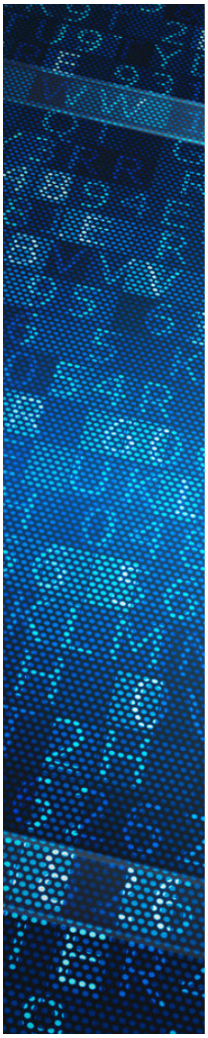
From the variables presented in the dropdown, select the relevant criminal variable(s) that are most important to you when trying to identify potentially malicious insider threats.

	Analytic Variable 1	Analytic Variable 2
Criminal		

10. List any additional variable(s) you think will support the identification of potentially malicious insiders cyber threats, that are not listed above.

* 11. From all the variables identified in Q1-Q9, please think of issues related to insider threat detection and list only the top 5 variables important to you when identifying potentially malicious insider threats

	Analytic Variables
Variable 1	
Variable 2	
Variable 3	
Variable 4	
Variable 5	



* 12. From the 5 variables you identified in Q11, please think of issues related to insider threat detection and rank the variables from most important to least important (1 – Most Important and 5 - Least Important).

- Variable 1
- Variable 2
- Variable 3
- Variable 4
- Variable 5

13. What are the positives of using the 5 variables you listed above when identifying potentially malicious insiders cyber threats?

14. Are there any negatives associated with using the 5 variables you listed above when identifying potentially malicious insiders cyber threats?

15. What is your job function?

- (A) Cybersecurity Engineer
- (B) Cybersecurity Analyst
- (C) Information Security Analyst
- (D) Information Security Engineer
- (E) Network Security Engineer
- (F) Information Technology Security Analyst
- (G) Information Security Manager
- (H) Information Assurance Engineer
- (I) Information Technology Auditor
- (J) Cybersecurity Administrator
- (K) Cybersecurity Consultant
- (L) Cybersecurity Architect
- Other (please specify)

16. Which describes your current employer?

- (A) Academia
- (B) Federal government employee
- (C) Private sector company
- (D) State government employee
- Other (please specify)

17. How long have you been with your current organization?

- (A) Under 1 year
- (B) 1 – 5 years
- (C) 6 – 10 years
- (D) 11 – 15 years
- (E) 16 – 20 years
- (F) 21 – 25 years
- (G) 26 – 30 years
- (H) Over 30 years

18. What is your highest level of education?

- (A) High school diploma
- (B) 2-year college (Associates degree)
- (C) 4-year college (Bachelors degree)
- (D) Graduate degree
- (E) Doctorate
- Other (please specify)

19. Do you currently hold any cybersecurity certifications, if so how many do you possess?

- (A) 0
- (B) 1
- (C) 2
- (D) 3
- (E) 4
- (F) 5 or more

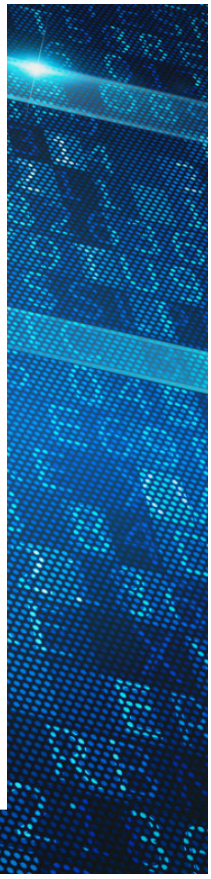
20. What is your age?

- (A) Under 20
- (B) 20-29
- (C) 30-39
- (D) 40-49
- (E) 50-59
- (F) Over 60

21. What is your gender?

- (A) Female
- (B) Male

Submit



Appendix F

Qualitative Survey Instrument 2 (FINAL): Email to SMEs

Google Forms

Input Requested from Cybersecurity Experts

Identified as an expert in the field, your assistance is needed to distinguish the most valid visualization techniques to present the top critical cyber visualization variables. The purpose of this survey is to identify the visualization techniques best suited to present the cyber visualization variables previously identified by SMEs in Survey Instrument 1. The goal is to identify visualization techniques that may aid in identifying potentially malicious insiders cyber threats.

I am a Ph.D. candidate in Information Systems at the College of Engineering and Computing, Nova Southeastern University, working under the supervision of Professor Yair Levy (levyy@nova.edu), and a member of his Levy CyLab (http://CyLab.nova.edu/). If you have any questions regarding this research, please do not hesitate to contact me or Dr. Levy via email.

Please send this invitation to participate in the survey to friends and colleagues with cybersecurity or visualization expertise. Thank you in advance for your time and consideration. I appreciate your assistance and contribution to this research study.

Regards,
Karla A. Clarke, Ph.D. Candidate

Qualitative Survey Instrument 2

Instrument for SME Identification of Visualization Technique for Cyber Variables

The purpose of this survey is to identify the visualization techniques best suited to present the cyber visualization variables as identified from Qualitative Survey Instrument 1. These variables may aid in efficiently identifying potentially malicious insiders cyber threats.

Based on data collection and analysis from Survey Instrument 1, six critical cyber visualization variables were identified. These variables were denoted as vital for the identification of potentially malicious insider cyber threats. Of the six variables identified by SMEs, two variables were tied for the number one spot as most critical. The identified variables are:

1. Workplace Satisfaction (#1)
2. Change in Violation Patterns (#1)
3. Audit Log Modification
4. Changes in Data Access Patterns
5. Data Exfiltration
6. Privilege Change

For the six variables identified above, this survey aims to identify visualization techniques that are most valid to present these top six critical cyber visualization variables. These variables are to be used when detecting potentially malicious cyber insider activities. This survey also aims to identify the visualization techniques that are most valid to present complex cyber data correlations relevant to the identified critical cyber visualization variables.

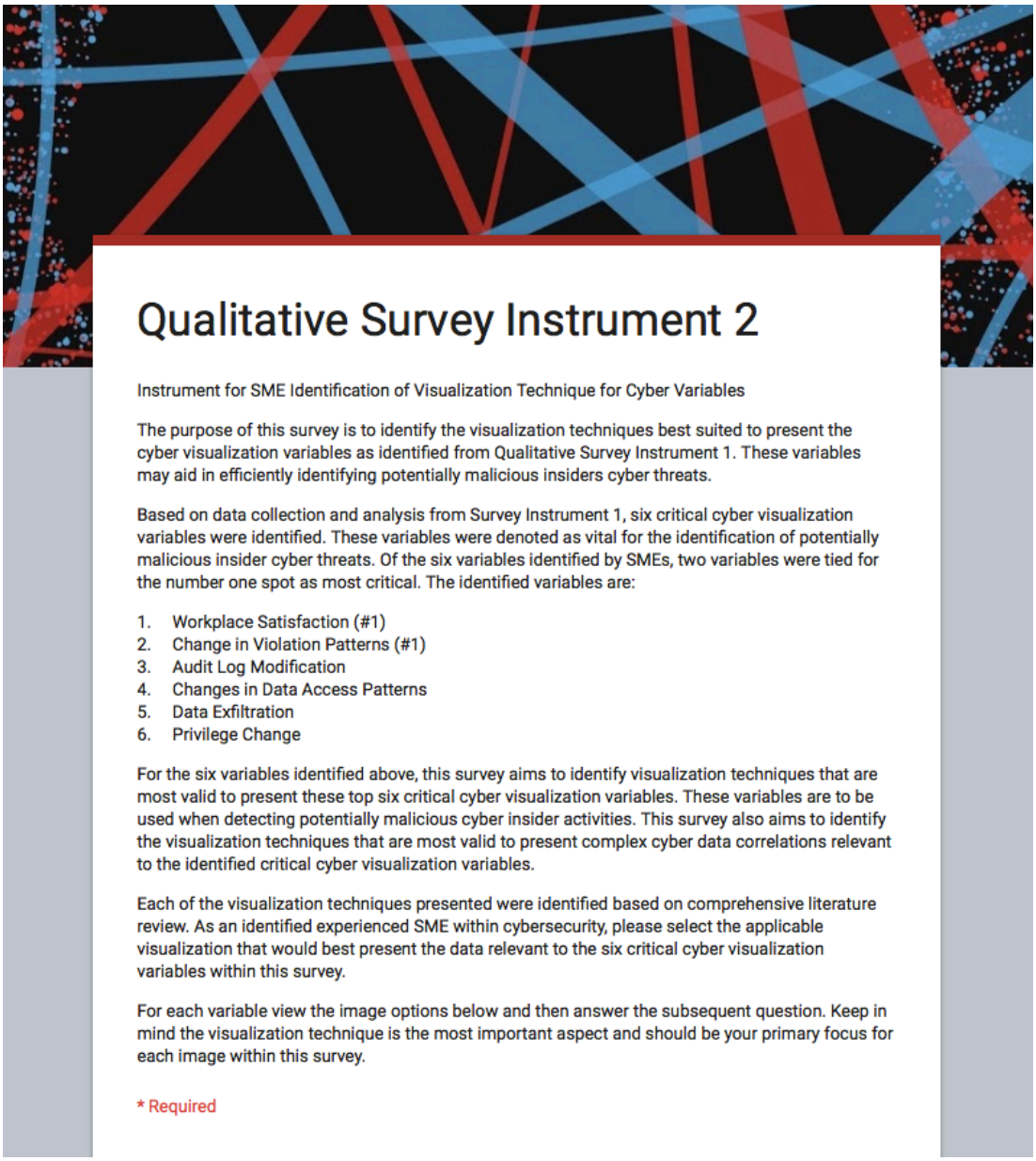
Each of the visualization techniques presented were identified based on comprehensive literature review. As an identified experienced SME within cybersecurity, please select the applicable visualization that would best present the data relevant to the six critical cyber visualization variables within this survey.

For each variable view the image options below and then answer the subsequent question. Keep in mind the visualization technique is the most important aspect and should be your primary focus for each image within this survey.

FILL OUT FORM

Appendix G

Qualitative Survey Instrument 2 (FINAL): Instrument for SME Identification of Visualization Technique for Cyber Variables



Qualitative Survey Instrument 2

Instrument for SME Identification of Visualization Technique for Cyber Variables

The purpose of this survey is to identify the visualization techniques best suited to present the cyber visualization variables as identified from Qualitative Survey Instrument 1. These variables may aid in efficiently identifying potentially malicious insiders cyber threats.

Based on data collection and analysis from Survey Instrument 1, six critical cyber visualization variables were identified. These variables were denoted as vital for the identification of potentially malicious insider cyber threats. Of the six variables identified by SMEs, two variables were tied for the number one spot as most critical. The identified variables are:

1. Workplace Satisfaction (#1)
2. Change in Violation Patterns (#1)
3. Audit Log Modification
4. Changes in Data Access Patterns
5. Data Exfiltration
6. Privilege Change

For the six variables identified above, this survey aims to identify visualization techniques that are most valid to present these top six critical cyber visualization variables. These variables are to be used when detecting potentially malicious cyber insider activities. This survey also aims to identify the visualization techniques that are most valid to present complex cyber data correlations relevant to the identified critical cyber visualization variables.

Each of the visualization techniques presented were identified based on comprehensive literature review. As an identified experienced SME within cybersecurity, please select the applicable visualization that would best present the data relevant to the six critical cyber visualization variables within this survey.

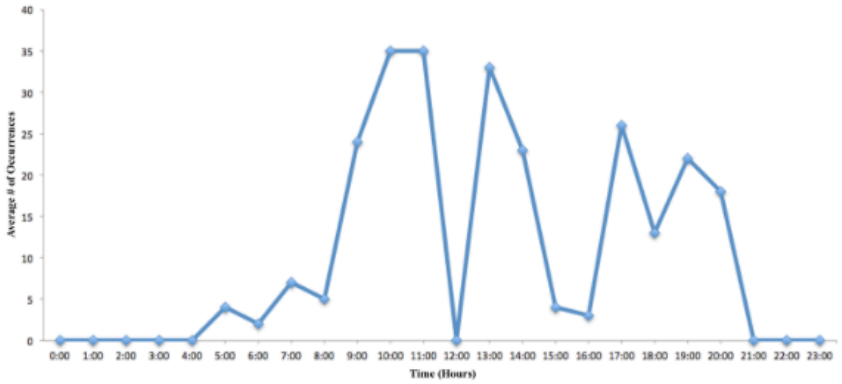
For each variable view the image options below and then answer the subsequent question. Keep in mind the visualization technique is the most important aspect and should be your primary focus for each image within this survey.

* Required

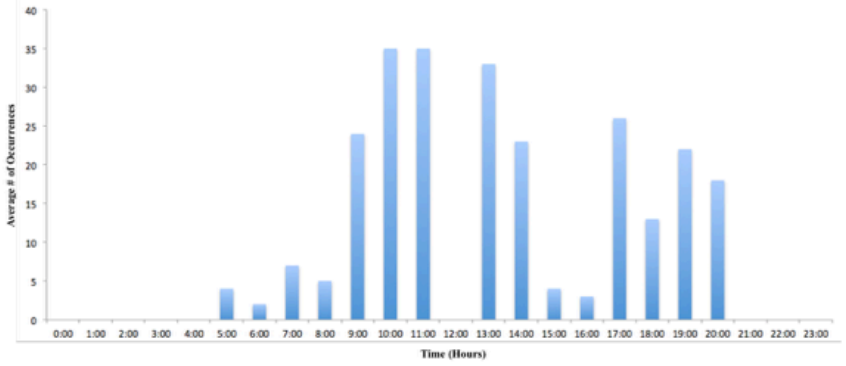
Analytic Variable 1: Workplace Satisfaction

Analytic Variable 1: Workplace Satisfaction - Since this variable does not generally have data available within a Security Information and Event Management (SIEM) system. Within literature this data may be identified using dictionary languages with theme-specific dictionaries that contain a list of words that have been validated to be associated with constructs. These constructs include wellbeing, engagement, positive and negative emotion, power etc. For instance, in the future this data may be collected from employee email, social media, or instant messenger content (Shami, Muller, Pal, Masli, & Geyer, 2015).

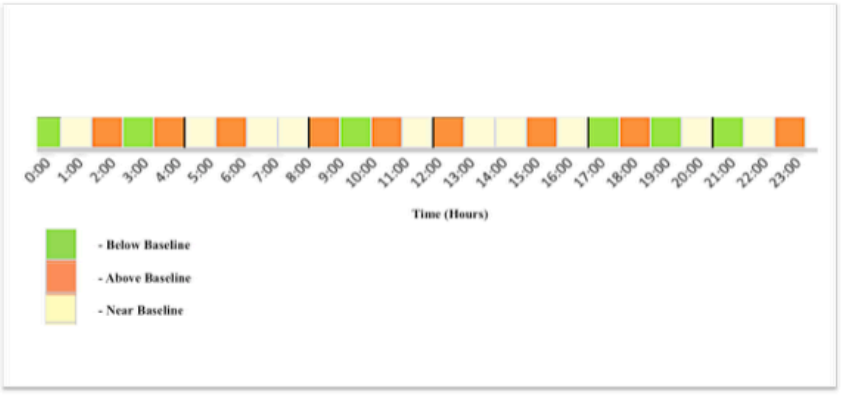
Analytic Variable 1: Workplace Satisfaction (Option 1)



Analytic Variable 1: Workplace Satisfaction (Option 2)



Analytic Variable 1: Workplace Satisfaction (Option 3)



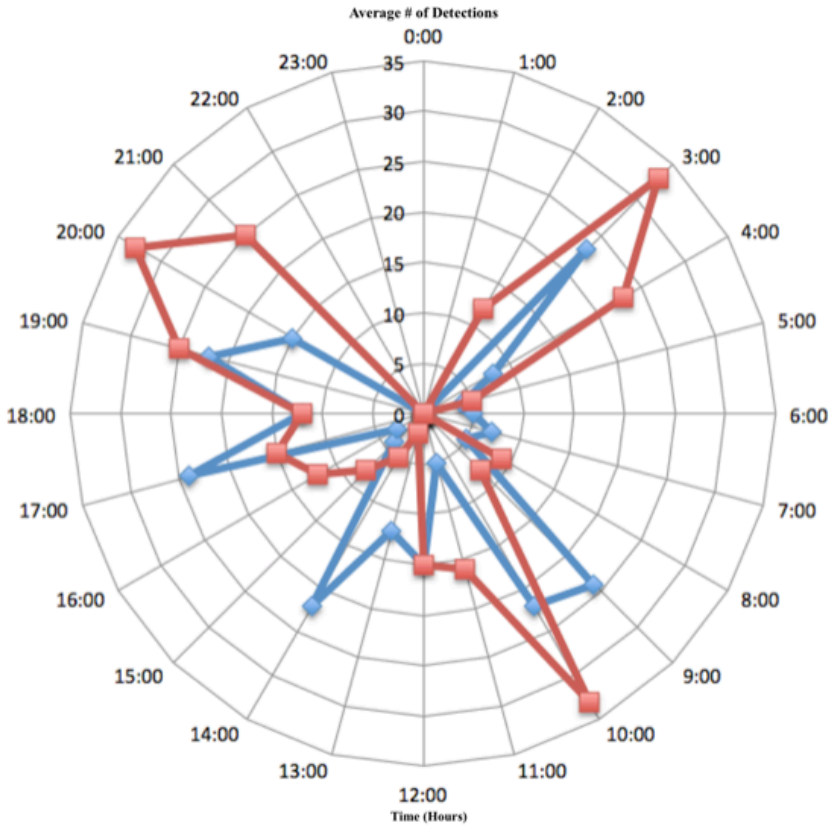
From the 3 options presented above select the visualization technique most relevant to displaying data related to workplace satisfaction for a potentially malicious insider. Based on an average number of occurrences within 24 hours.

- ◯ Option 1: Line Graph
- ◯ Option 2: Bar Graph
- ◯ Option 3: Calendar View

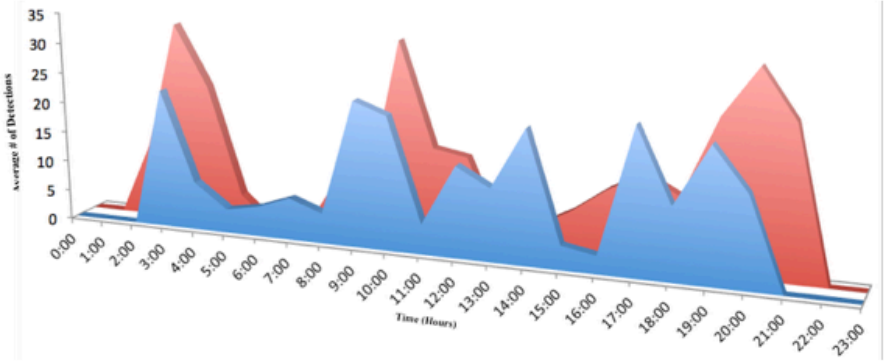
Analytic Variable 2: Change in Violation Patterns

Analytic Variable 2: Change in Violation Patterns - For the purpose of this research a change in violation pattern will refer to a deviation in an initially identified pattern of user behaviors that result in violations. Violation patterns are generally identified using logs. Based on a set threshold of detections, a pattern of user behavior may be identified. An organization may identify specific triggers for detected events. For the purpose of this research study examples of events which may be used for the development of violation patterns include: software installation, managing system services, as well as successful and failed login attempts (Malec, Piwowar, Kozakiewicz, & Lasota, 2015).

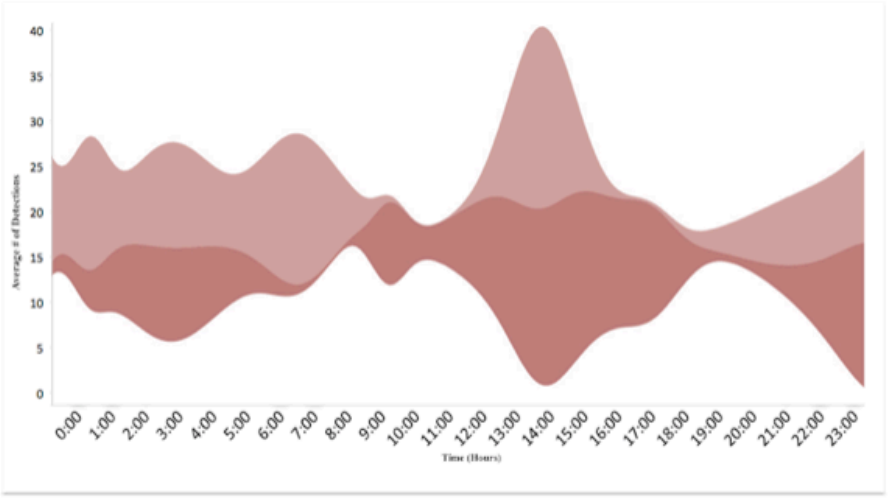
Analytic Variable 2: Change in Violation Patterns (Option 1)



Analytic Variable 2: Change in Violation Patterns (Option 2)



Analytic Variable 2: Change in Violation Patterns (Option 3)



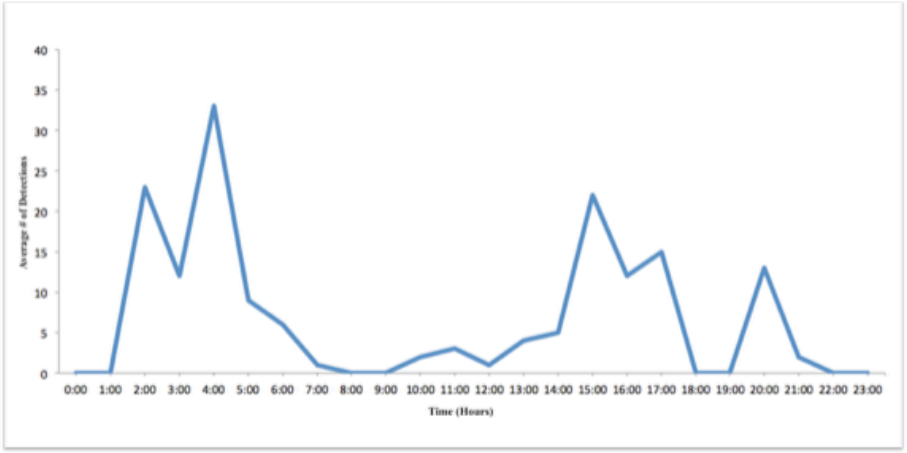
Select the visualization technique presented in the 3 options above, most relevant to displaying data related to changes in violation patterns for a potentially malicious insider. In the image for option 1 and 2 blue represents the users initially identified violation pattern and red represents the newly identified violation pattern. For option 3 the dark red blue represents the users initially identified violation pattern and light red represents the newly identified violation pattern. Based on an average number of detections within 24 hours.

- ◯ Option 1: Radar Plot
- ◯ Option 2: Area Chart
- ◯ Option 3: Streamgraph

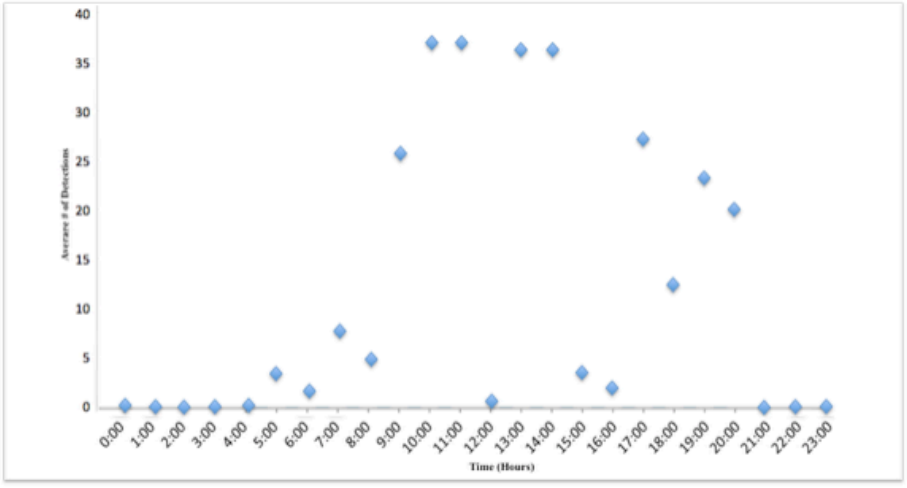
Analytic Variable 3: Audit Log Modification

Analytic Variable 3: Audit Log Modification - Log files contain messages emitted from several modules within a system, these logs may include information such as device status and error conditions and also about the various tasks within the system such as program names, execution path, including function names and parameters, and the task completion status (Basak & Nagesh, 2016). Small modifications to these logs may allow for successful evasion of signature methods when identifying a potentially malicious insider (Berlin, Slater, & Saxe, 2015).

Analytic Variable 3: Audit Log Modification (Option 1)



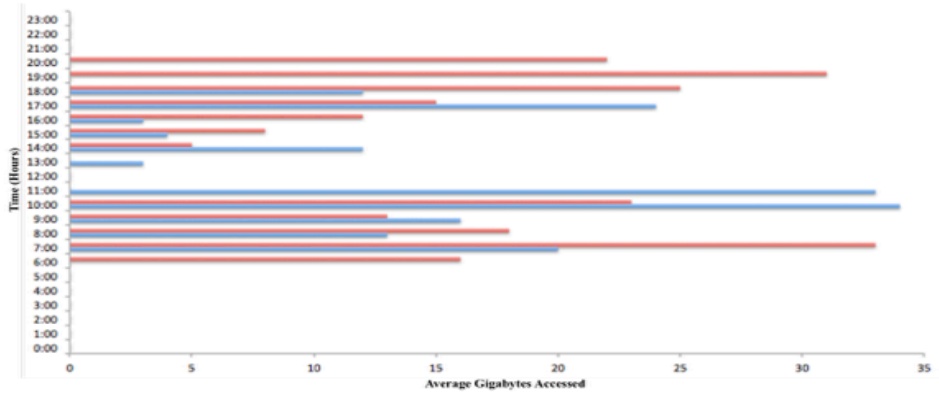
Analytic Variable 3: Audit Log Modification (Option 2)



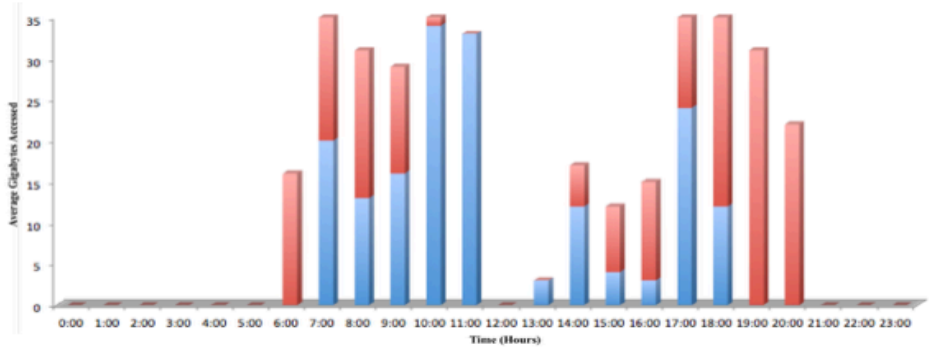
Analytic Variable 4: Changes in Data Access Patterns

Analytic Variable 4: Changes in Data Access Patterns - For the purpose of this research a change in data access pattern will refer to a deviation in an initially identified access pattern of a potentially malicious user. Users within similar roles may be expected to have similar data access patterns, by identifying patterns within roles an organization my develop a 'fingerprint' of historical access patterns (Menon, Jiang, Kim, Vaidya, & Ohno-Machado, 2013). In the event of deviations from these patterns a potentially malicious insider may be more efficiently identified.

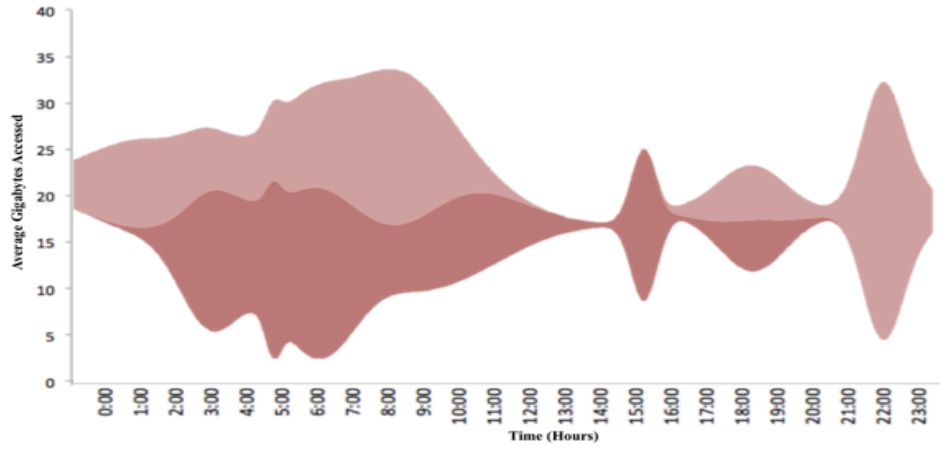
Analytic Variable 4: Changes in Data Access Patterns (Option 1)



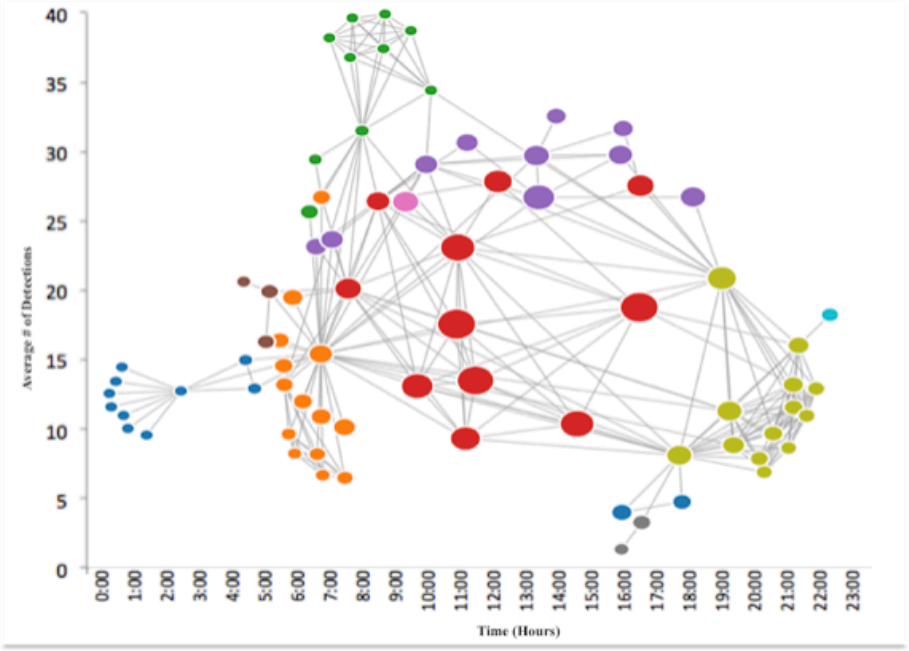
Analytic Variable 4: Changes in Data Access Patterns (Option 2)



Analytic Variable 4: Changes in Data Access Patterns (Option 3)



Analytic Variable 3: Audit Log Modification (Option 3)



From the 3 options presented above select the visualization technique most relevant to displaying data related to audit log modifications by a potentially malicious insider. Based on an average number of detections within 24 hours at the time in which the log was modified.

- ◯ Option 1: Line Graph
- ◯ Option 2: Bar Graph
- ◯ Option 3: Fisheye Distortion

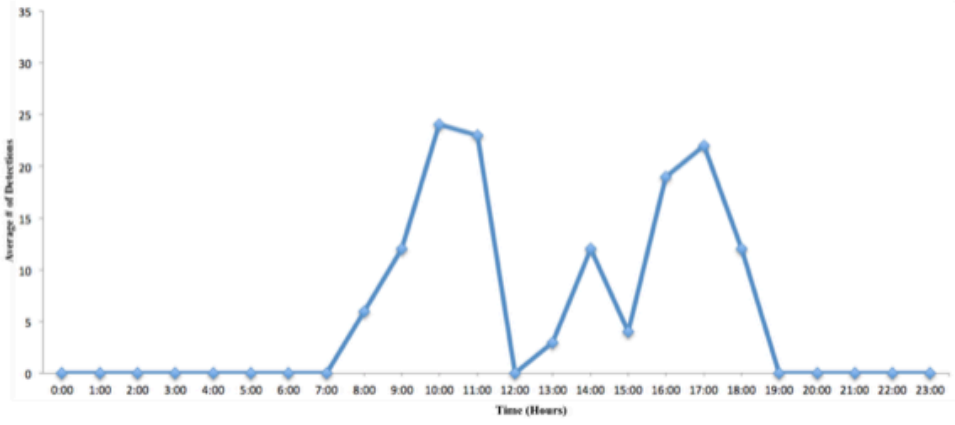
From the 3 options presented above select the visualization technique most relevant to displaying data related to changes in violation patterns. In the image for option 1 and 2 blue represents the users initially identified violation pattern and red represents the newly identified violation pattern. For option 3 the dark red blue represents the users initially identified violation pattern and light red represents the newly identified violation pattern. Based on the average number of detections within 24 hours.

- ◯ Option 1: Stacked Bar Graph
- ◯ Option 2: Stacked Column Graph
- ◯ Option 3: Streamgraph

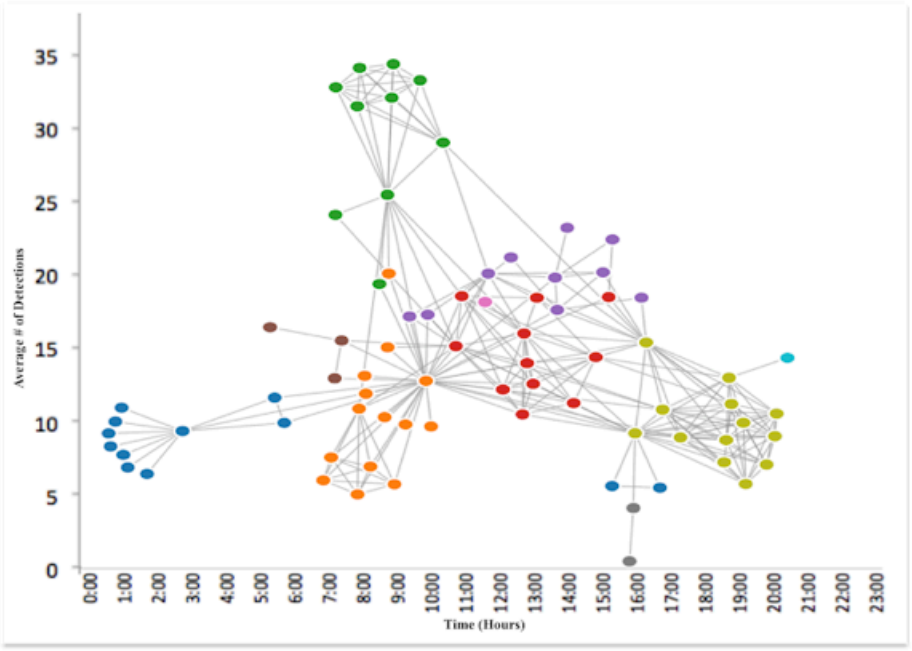
Analytic Variable 5: Data Exfiltration

Analytic Variable 5: Data Exfiltration - For the purpose of this research study data exfiltration refers to transferring sensitive data outside the protected domain. This is generally performed to steal intelligence, identities, or credentials. By identifying patterns in user network traffic using information such as: source IP, destination IP, port numbers, flow durations, and payload size an organization can identify data flow levels or thresholds (Wang, Yang, & Chen, 2015). These thresholds may be use to efficiently identify data exfiltration by a potentially malicious insider.

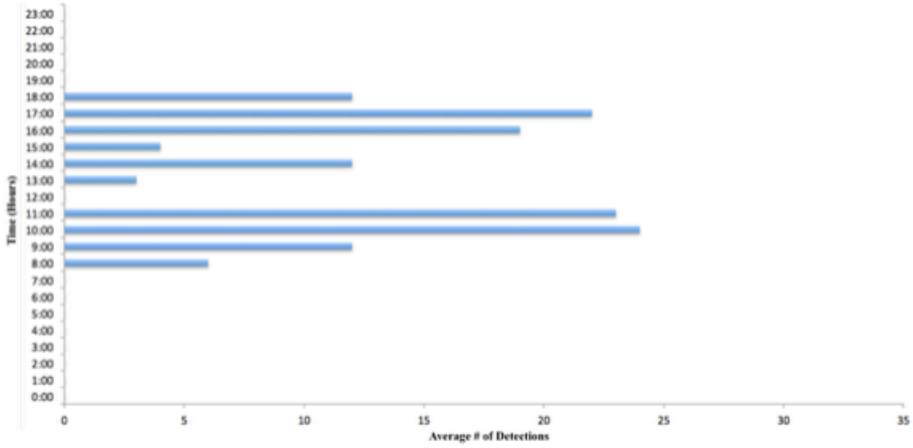
Analytic Variable 5: Data Exfiltration (Option 1)



Analytic Variable 5: Data Exfiltration (Option 2)



Analytic Variable 5: Data Exfiltration (Option 3)



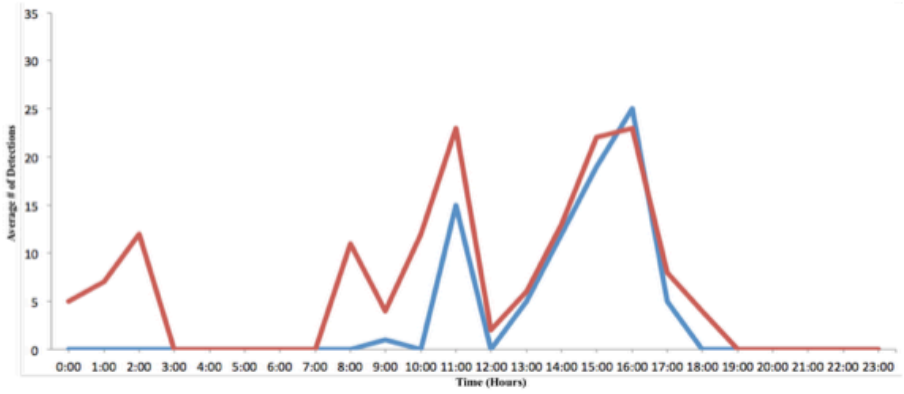
From the 3 options presented above select the visualization technique most relevant to displaying data related to data exfiltration by a potentially malicious insider. Based on the average number of detections within 24 hours.

- ◯ Option 1: Line Graph
- ◯ Option 2: Fisheye Distortion
- ◯ Option 3: Column Graph

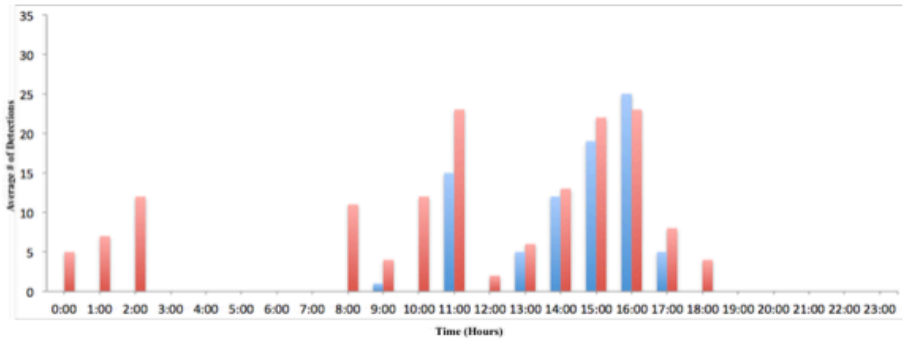
Analytic Variable 6: Privilege Change

Analytic Variable 6: Privilege Change - With principles of least privilege and privilege separation access control is deemed as a vital to protecting confidentiality and integrity of data. Failing to properly separate privileges leave applications vulnerable to attacks such as privilege escalation (Hsu, Hoffman, Eugster, & Payer, 2016). For this research study privilege change refers to privilege escalations, for instance, the elevation of a standard user account to a user account with administrator privileges.

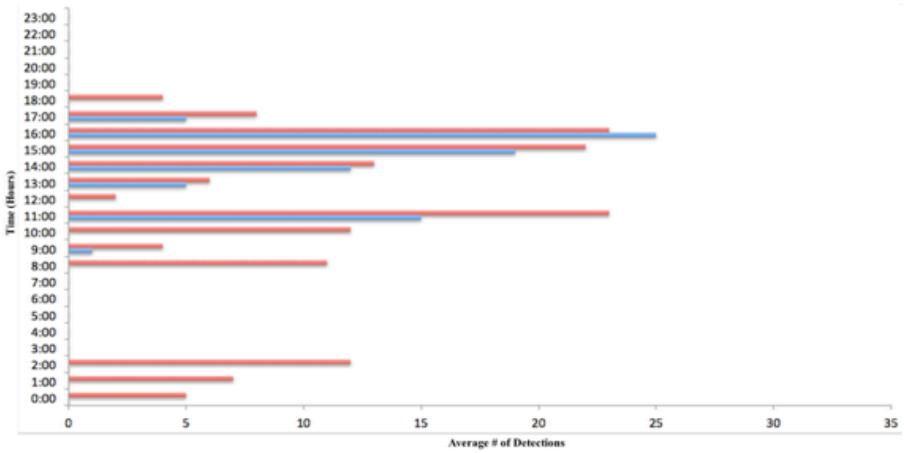
Analytic Variable 6: Privilege Change (Option 1)



Analytic Variable 6: Privilege Change (Option 2)



Analytic Variable 6: Privilege Change (Option 3)



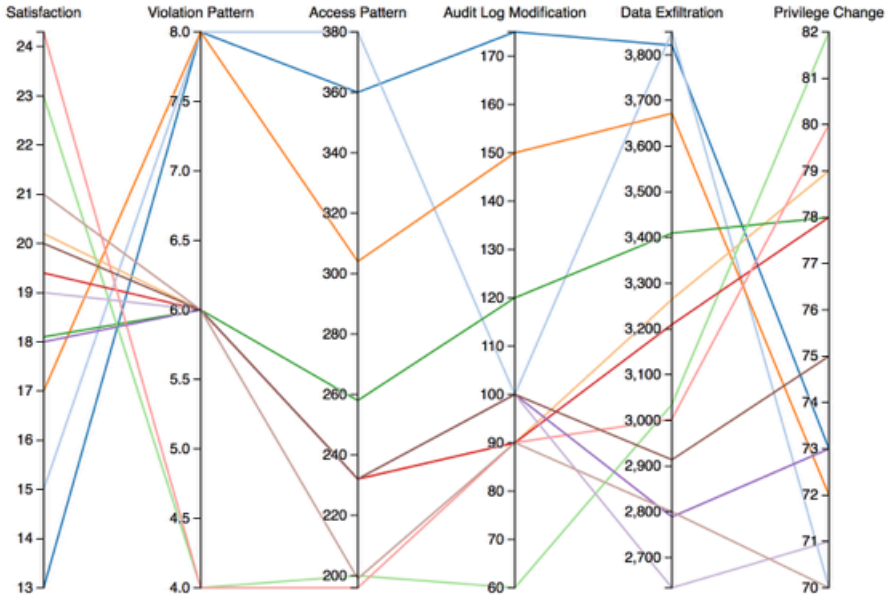
From the 3 options presented above select the visualization technique most relevant to displaying data related to privilege change by a potentially malicious insider. In the images blue represents the users initially identified violation pattern and red represents the newly identified violation pattern. Based on an average number of detections within 24 hours.

- ◯ Option 1: Line Graph
- ◯ Option 2: Stacked Bars
- ◯ Option 3: Stacked Columns

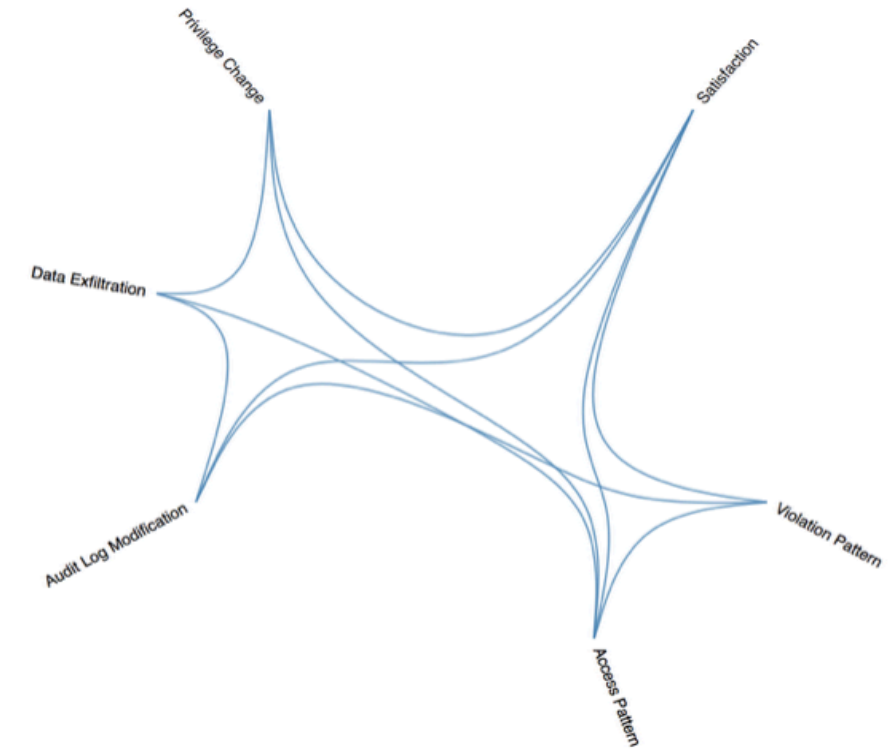
Visualization of Complex Cyber Data Correlations

Visualization of Complex Cyber Data Correlations - Potentially malicious insiders are unpredictable and their malicious along side non-malicious actions add complexity to the identification of insider threats. This necessitates the careful analysis of network, system and user parameters often correlated with other user identity data (Gamachchi & Boztas, 2017). As a result anomaly detection techniques often entail correlations of multiple data sets resulting in complex data correlations.Within this research study a complex data correlation will refer to identifying linear or non-linear relationships between two or more data variables (Patcha & Park, 2007).

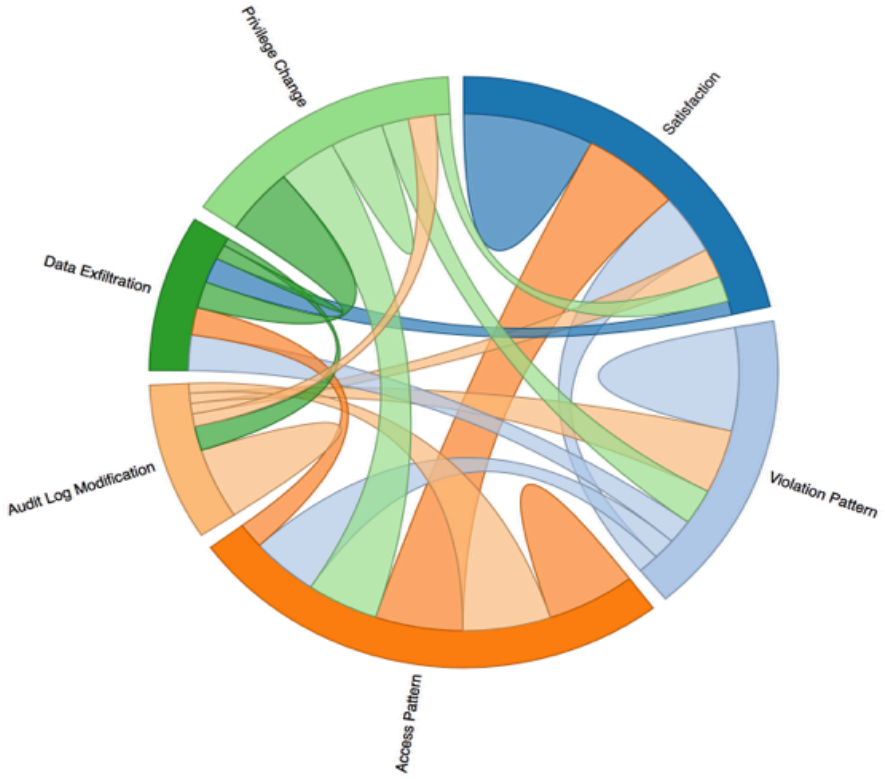
Visualization of Complex Cyber Data Correlations (Option 1)



Visualization of Complex Cyber Data Correlations (Option 2)



Visualization of Complex Cyber Data Correlations (Option 3)



From the top 6 analytic variables identified select the visualization technique from the 3 options presented above that are most relevant to displaying two or more data variables. Based on an average number of detections within 24 hours for all the identified critical variables.

- ◯ Option 1: Parallel Coordinates
- ◯ Option 2: Hierarchical Bundling
- ◯ Option 3: Chord Diagram

Demographic Information

What is your job function? *

- ◯ Cybersecurity Engineer
- ◯ Cybersecurity Analyst
- ◯ Information Security Analyst
- ◯ Network Security Engineer
- ◯ Information Technology Security Analyst
- ◯ Information Security Manager
- ◯ Information Technology Auditor
- ◯ Cybersecurity Administrator
- ◯ Cybersecurity Consultant
- ◯ Cybersecurity Architect
- ◯ Other: _____

What describes your current employer? *

- ◯ Academia
- ◯ Federal government employee
- ◯ State government employee
- ◯ Other: _____

How long have you been with your current employer? *

- ◯ Under 1 year
- ◯ 1-5 years
- ◯ 6-10 years
- ◯ 11-15 years
- ◯ 16-20 years
- ◯ 21-25 years
- ◯ Over 30 years

What is your highest level of education? *

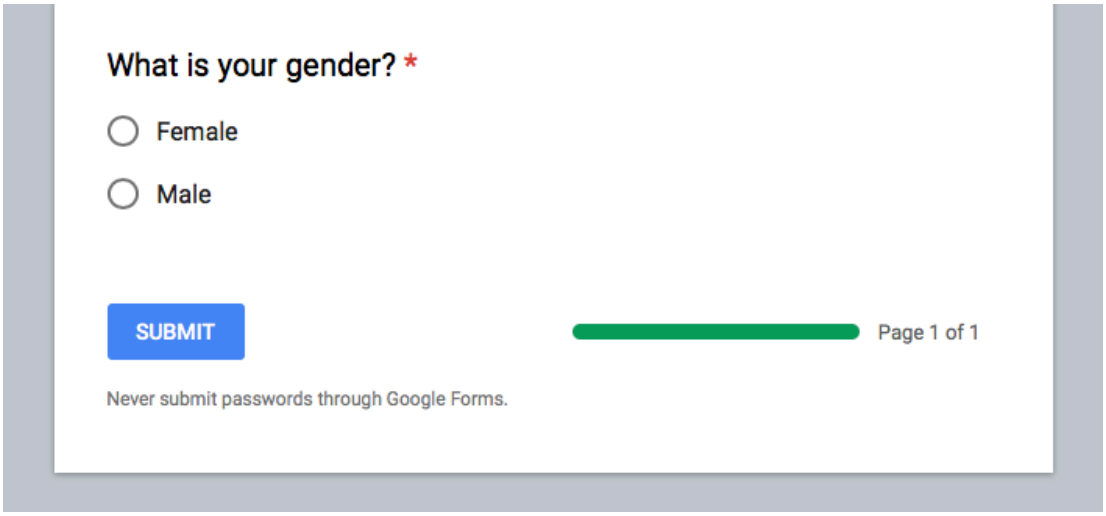
- ◯ High school diploma
- ◯ 2-year college (Associates degree)
- ◯ 4-year college (Bachelors Degree)
- ◯ Graduate degree
- ◯ Doctorate
- ◯ Other: _____

Do you currently hold and cybersecurity certifications, if so how many do you possess? *

- ◯ 0
- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5 or more

What is your age? *

- ◯ Under 20
- ◯ 20-29
- ◯ 30-39
- ◯ 40-49
- ◯ 50-59
- ◯ Over 60


What is your gender? *
Female
Male
SUBMIT
Page 1 of 1
Never submit passwords through Google Forms.

Appendix H

Qualitative Survey Instrument 3 (FINAL): Email to SMEs

Google Forms

Experience the newly developed prototype QUICK.v™

Identified as an expert in the field, your assistance is needed to complete the final survey to validate the effectiveness (i.e. satisfaction & value/importance) of the QUICK.v™ prototype. The developed prototype is to be used when mitigating potentially malicious cyber insider threats. The purpose of this survey is to identify the effectiveness of the developed prototype QUICK.v™.

I am a Ph.D. candidate in Information Systems at the College of Engineering and Computing, Nova Southeastern University, working under the supervision of Professor Yair Levy (levyy@nova.edu), and a member of his Levy CyLab (http://CyLab.nova.edu/). If you have any questions regarding this research, please do not hesitate to contact me or Dr. Levy via email.

Send this invitation to participate in the survey to any friends and colleagues with cybersecurity or visualization expertise. Thank you in advance for your time and consideration. I appreciate your assistance and contribution to this research study.

Regards,
Karla A. Clarke, Ph.D. Candidate
Kc1127@mynsu.nova.edu

Instrument for Cybersecurity Analysts Effectiveness

Dear participants,

The purpose of this survey is to determine the perceived effectiveness of the cyber visualization variables presented and the visualization techniques used. The instrument depicted at the link below is the final version that was developed based on Subject Matter Experts' (SMEs) input.

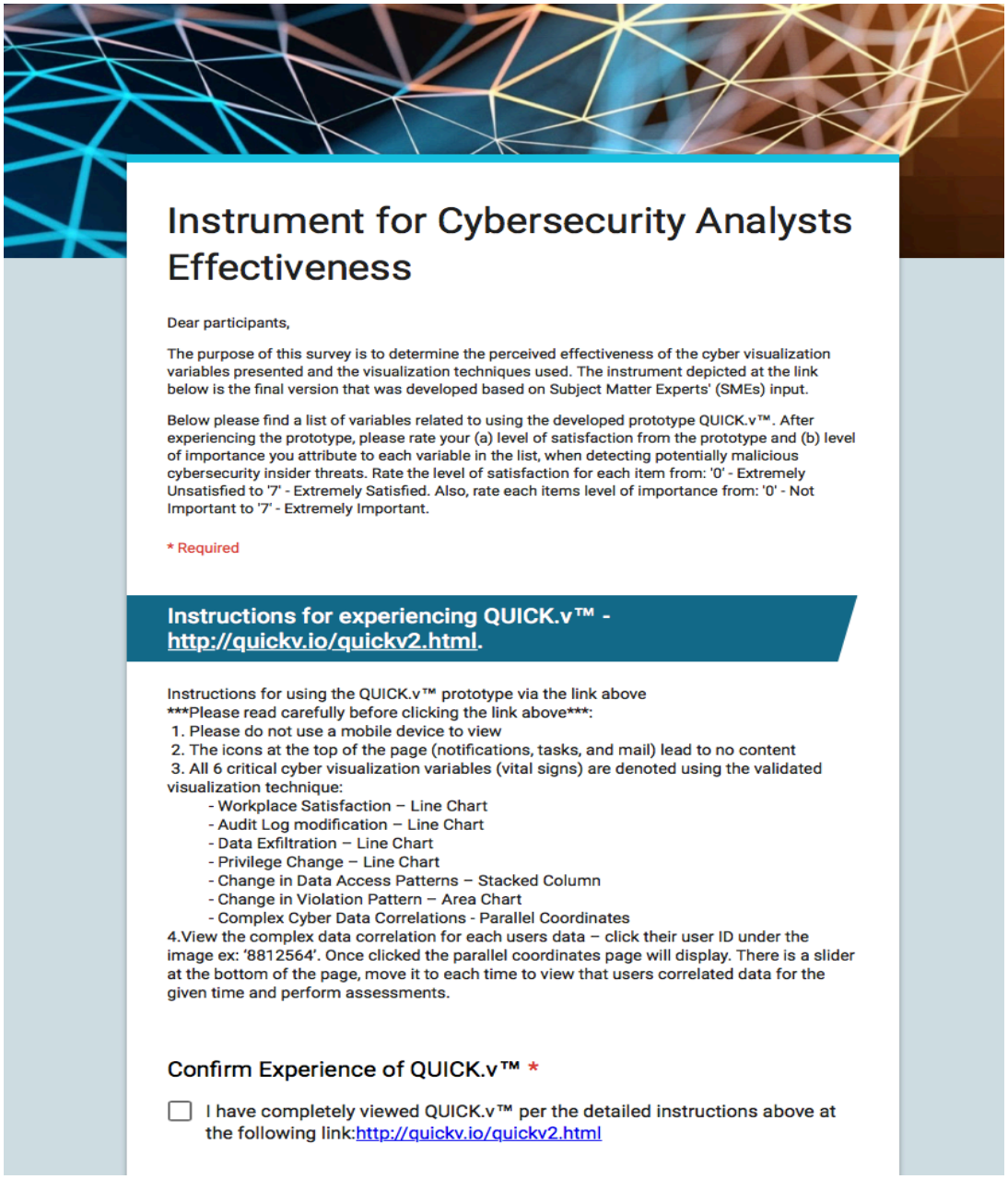
Below please find a list of variables related to using the developed prototype QUICK.v™. After experiencing the prototype, please rate your (a) level of satisfaction from the prototype and (b) level of importance you attribute to each variable in the list, when detecting potentially malicious cybersecurity insider threats. Rate the level of satisfaction for each item from: '0' - Extremely Unsatisfied to '7' - Extremely Satisfied. Also, rate each items level of importance from: '0' - Not Important to '7' - Extremely Important.

FILL OUT FORM

Create your own Google Form

Appendix I

Quantitative Survey Instrument 3 (FINAL): Instrument for Cybersecurity Analysts' Perceived Effectiveness of the Prototype



Instrument for Cybersecurity Analysts Effectiveness

Dear participants,

The purpose of this survey is to determine the perceived effectiveness of the cyber visualization variables presented and the visualization techniques used. The instrument depicted at the link below is the final version that was developed based on Subject Matter Experts' (SMEs) input.

Below please find a list of variables related to using the developed prototype QUICK.v™. After experiencing the prototype, please rate your (a) level of satisfaction from the prototype and (b) level of importance you attribute to each variable in the list, when detecting potentially malicious cybersecurity insider threats. Rate the level of satisfaction for each item from: '0' - Extremely Unsatisfied to '7' - Extremely Satisfied. Also, rate each items level of importance from: '0' - Not Important to '7' - Extremely Important.

* Required

Instructions for experiencing QUICK.v™ - http://quickv.io/quickv2.html.

Instructions for using the QUICK.v™ prototype via the link above
Please read carefully before clicking the link above:
1. Please do not use a mobile device to view
2. The icons at the top of the page (notifications, tasks, and mail) lead to no content
3. All 6 critical cyber visualization variables (vital signs) are denoted using the validated visualization technique:
   - Workplace Satisfaction – Line Chart
   - Audit Log modification – Line Chart
   - Data Exfiltration – Line Chart
   - Privilege Change – Line Chart
   - Change in Data Access Patterns – Stacked Column
   - Change in Violation Pattern – Area Chart
   - Complex Cyber Data Correlations - Parallel Coordinates
4. View the complex data correlation for each users data – click their user ID under the image ex: '8812564'. Once clicked the parallel coordinates page will display. There is a slider at the bottom of the page, move it to each time to view that users correlated data for the given time and perform assessments.

Confirm Experience of QUICK.v™ *

☐ I have completely viewed QUICK.v™ per the detailed instructions above at the following link:http://quickv.io/quickv2.html

A. The following items relate to the cybersecurity variables presented in QUICK.v™:

Variable 1: Workplace Satisfaction

Level of Satisfaction with "Variable 1: Workplace Satisfaction" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 1: Workplace Satisfaction" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Variable 2: Audit Log Modification

Level of Satisfaction with "Variable 2: Audit Log Modification" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 2: Audit Log Modification" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Variable 3: Data Exfiltration

Level of Satisfaction with "Variable 3: Data Exfiltration" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 3: Data Exfiltration" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Variable 4: Privilege Change

Level of Satisfaction with "Variable 4: Privilege Change" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 4: Privilege Change" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Variable 5: Changes in Data Access Patterns

Level of Satisfaction with "Variable 5: Changes in Data Access Patterns" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 5: Changes in Data Access Patterns" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Variable 6: Change in Violation Patterns

Level of Satisfaction with "Variable 6: Change in Violation Patterns" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Variable 6: Change in Violation Patterns" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Other Variable: Complex Cyber Data Correlations

Level of Satisfaction with "Other Variable: Complex Cyber Data Correlations" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Other Variable: Complex Cyber Data Correlations" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Type of Variables Presented

Level of Satisfaction with "Type of Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Type of Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Interest in Variables Presented

Level of Satisfaction with "Interest in Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Interest in Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Organization of Variables Presented

Level of Satisfaction with "Organization of Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Organization of Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Complexity Based on Variables Presented

Level of Satisfaction with "Complexity Based on Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Complexity Based on Variables Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Various Variables Were Integrated Well

Level of Satisfaction with "Various Variables Were Integrated Well" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Various Variables Were Integrated Well" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Relevance of Variables to Insider Threat Detection

Level of Satisfaction with "Relevance of Variables to Insider Threat Detection" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Relevance of Variables to Insider Threat Detection" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

B. The following items relate to the visualization methods used to present the cybersecurity variables:

Quality of Visualizations

Level of Satisfaction with "Quality of Visualizations" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Quality of Visualizations" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Organization of Visualizations Presented

Level of Satisfaction with "Organization of Visualizations Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Organization of Visualizations Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Consistency (accuracy) of Visualizations Presented

Level of Satisfaction with "Consistency (accuracy) of Visualizations Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Consistency (accuracy) of Visualizations Presented" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Ability to Quickly Decipher Potential Insider Threats

Level of Satisfaction with "Ability to Quickly Decipher Potential Insider Threats" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Ability to Quickly Decipher Potential Insider Threats" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Confidence Quickly Deciphering Potential Insider Threats

Level of Satisfaction with "Confidence Quickly Deciphering Potential Insider Threats" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

C. Global Questions:

Overall, how would you rate your level of satisfaction with QUICK.v™ when identifying potentially malicious cyber insiders?

Level of Satisfaction *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Overall, how important would QUICK.v™ be to you when identifying potentially malicious cyber insiders?

Level of Importance *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Demographic Information

What is your current job function? *

- ○ Cybersecurity Engineer
- ○ Cybersecurity Analyst
- ○ Information Security Analyst
- ○ Network Security Engineer
- ○ Information Technology Security Analyst
- ○ Information Security Manager
- ○ Information Technology Auditor
- ○ Cybersecurity Administrator
- ○ Cybersecurity Consultant
- ○ Cybersecurity Architect
- ○ Other: _____

Level of Importance of "Confidence Quickly Deciphering Potential Insider Threats" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Ability to Make Actionable Decisions Based on Information Depicted

Level of Satisfaction with "Ability to Make Actionable Decisions Based on Information Depicted" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Ability to Make Actionable Decisions Based on Information Depicted" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Ease of Use of Information Depicted

Level of Satisfaction with "Ease of Use of Information Depicted" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

Level of Importance of "Ease of Use of Information Depicted" *

1	2	3	4	5	6	7
○	○	○	○	○	○	○

What describes your current employer? *

- ◯ Academia
- ◯ Federal government employee
- ◯ State government employee
- ◯ Other: _____

How long have you been with your current employer? *

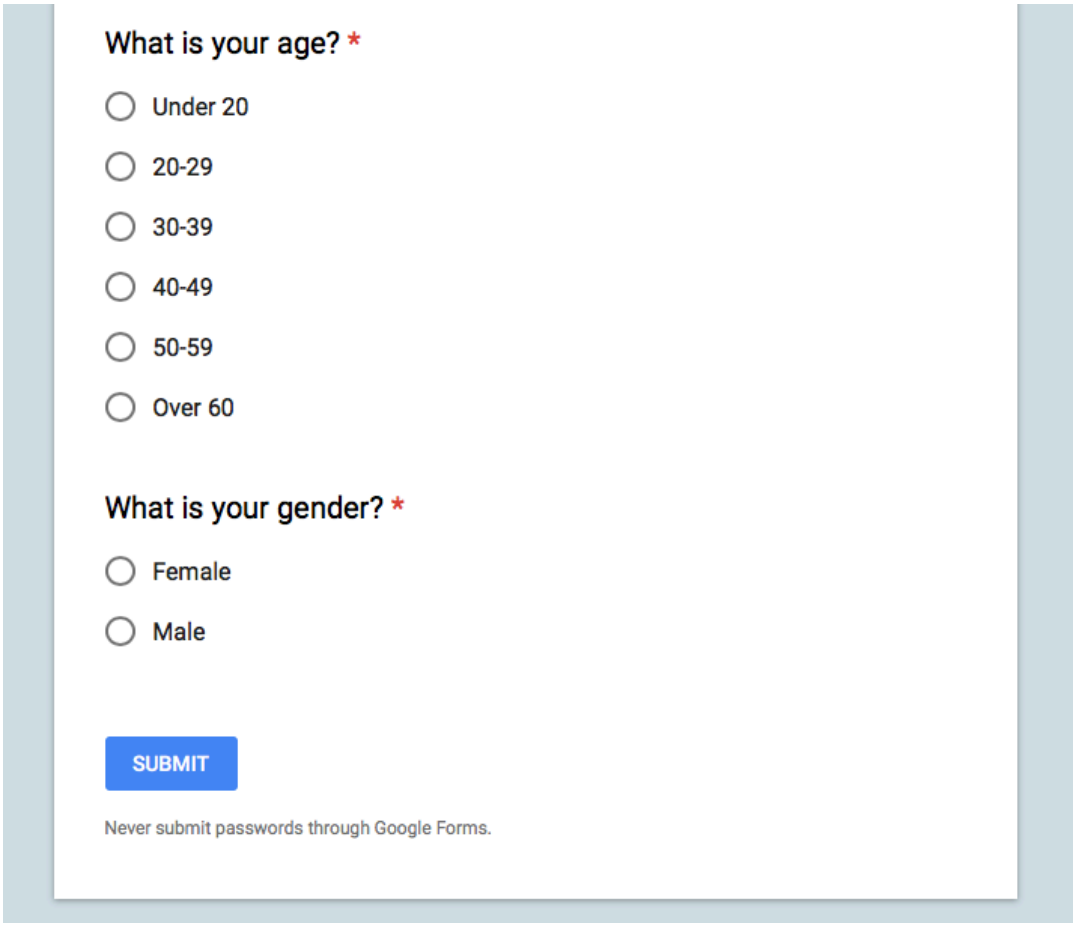
- ◯ Under 1 year
- ◯ 1-5 years
- ◯ 6-10 years
- ◯ 11-15 years
- ◯ 16-20 years
- ◯ 21-25 years
- ◯ Over 30 years

What is your highest level of education? *

- ◯ High school diploma
- ◯ 2-year college (Associates degree)
- ◯ 4-year college (Bachelors Degree)
- ◯ Graduate degree
- ◯ Doctorate
- ◯ Other: _____

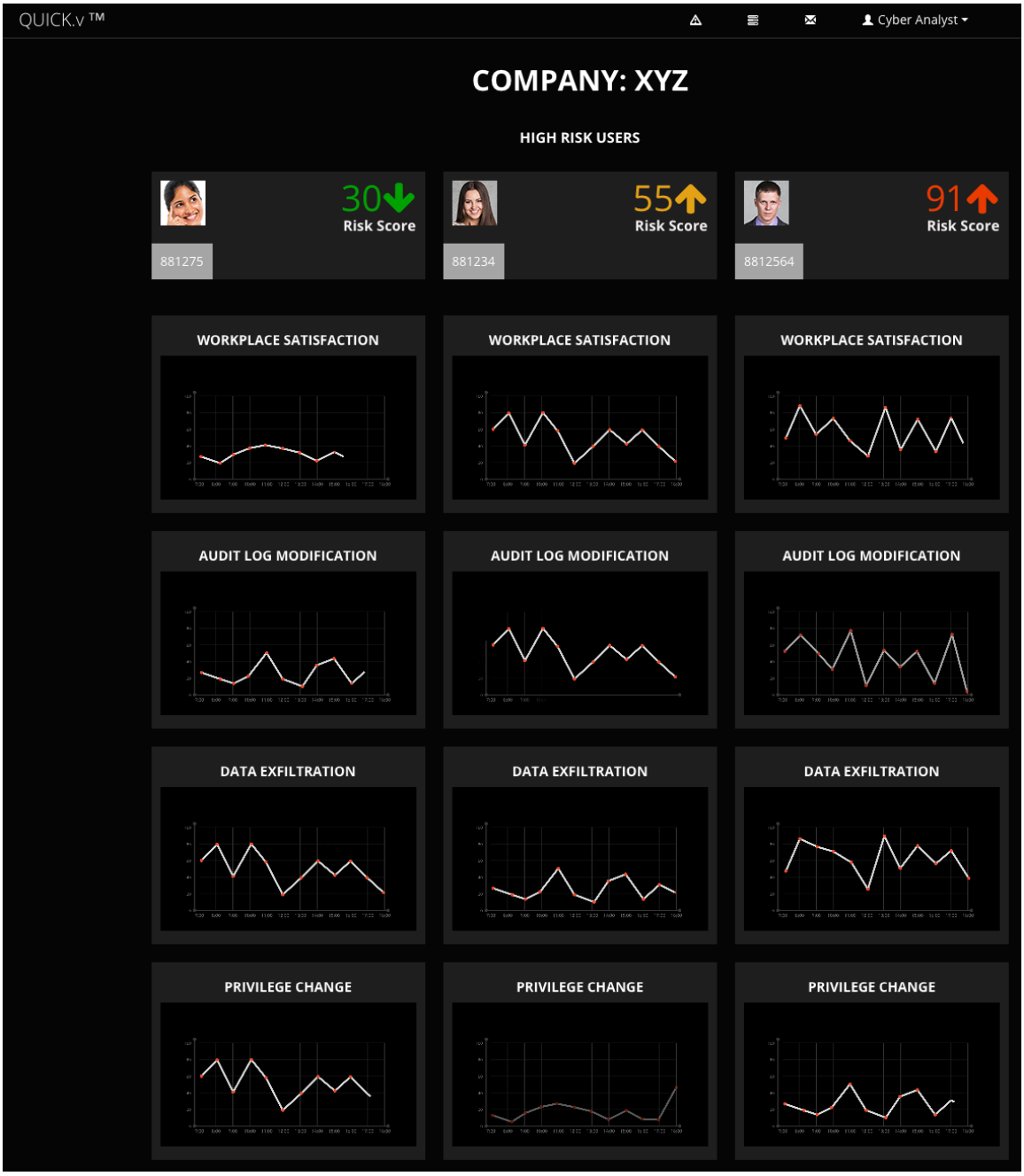
Do you currently hold and cybersecurity certifications, if so how many do you possess? *

- ◯ 0
- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5 or more

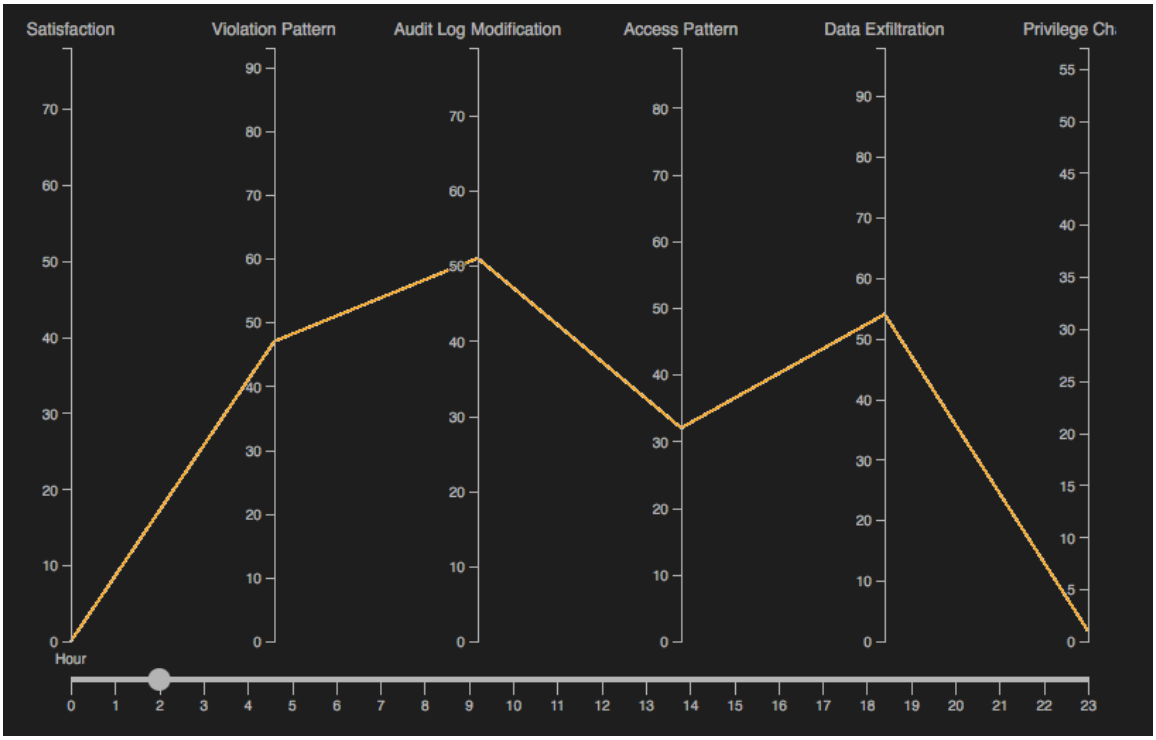

What is your age? *
Under 20
20-29
30-39
40-49
50-59
Over 60
What is your gender? *
Female
Male
SUBMIT
Never submit passwords through Google Forms.

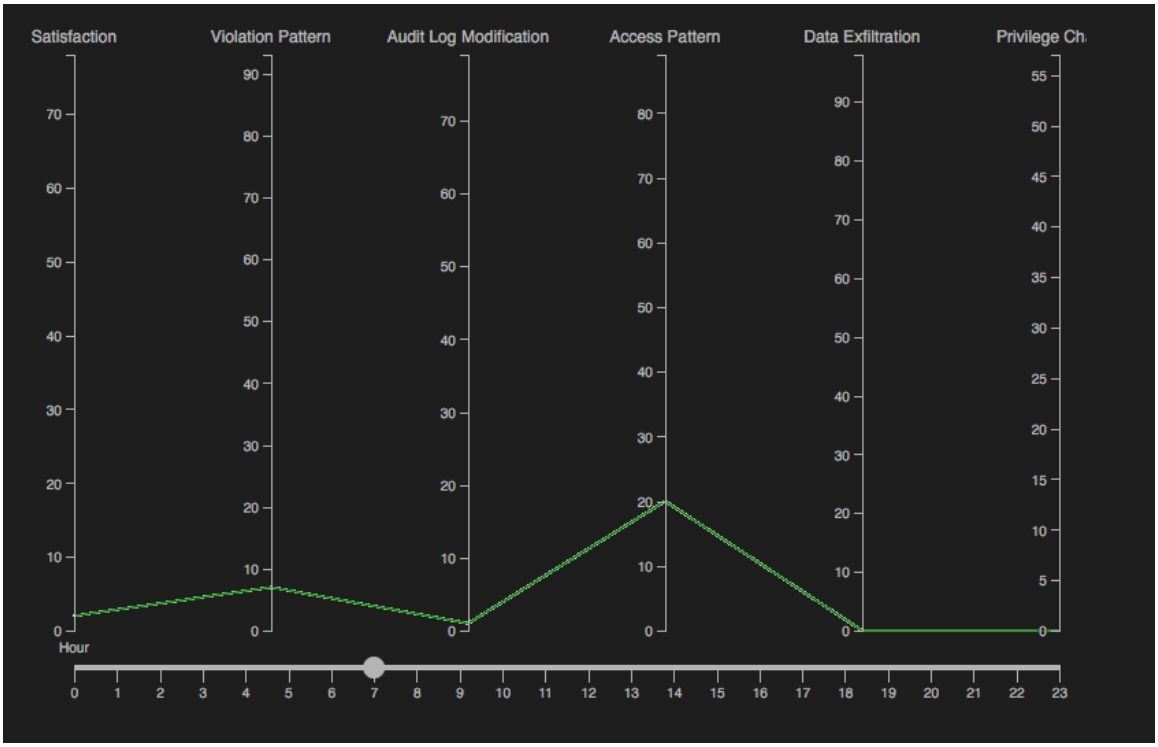
Appendix J

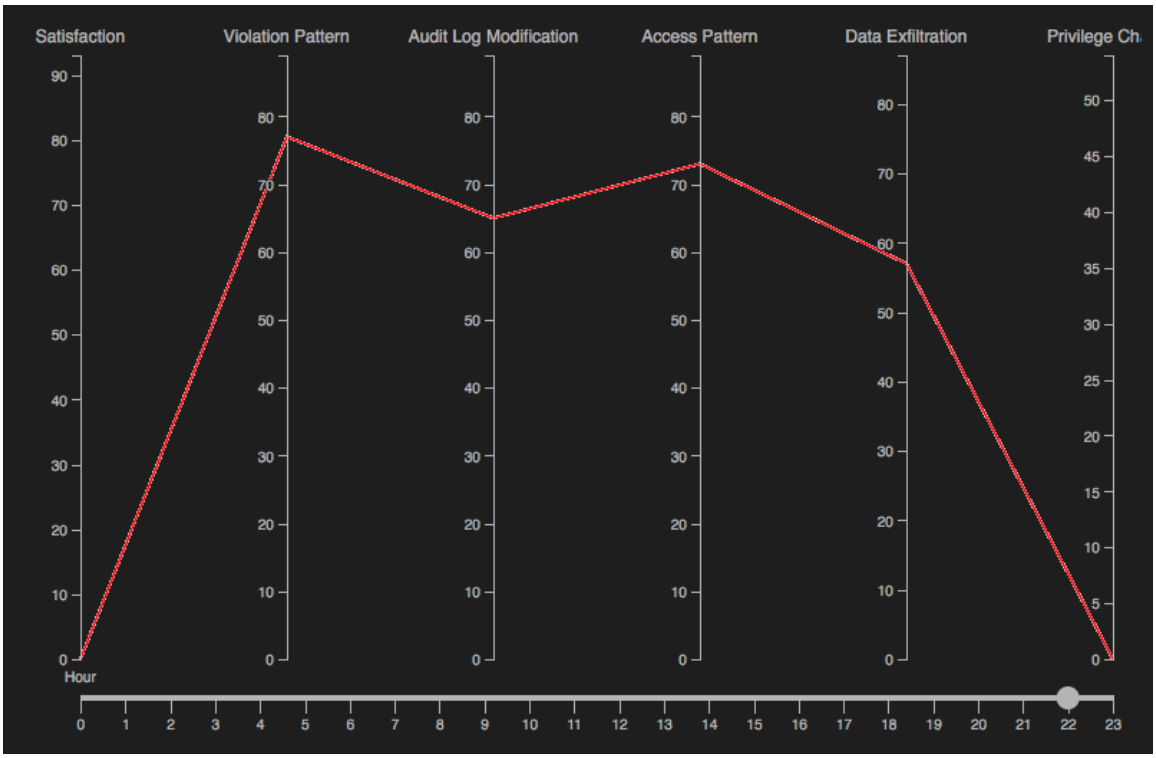
Developed Prototype QUICK.v™ (FINAL)




CHANGE IN DATA ACCESS PATTERN
CHANGE IN DATA ACCESS PATTERN
CHANGE IN DATA ACCESS PATTERN
CHANGE IN VIOLATION PATTERN
CHANGE IN VIOLATION PATTERN
CHANGE IN VIOLATION PATTERN


Satisfaction
Violation Pattern
Audit Log Modification
Access Pattern
Data Exfiltration
Privilege Ch
Hour
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23


Satisfaction
Violation Pattern
Audit Log Modification
Access Pattern
Data Exfiltration
Privilege Ch
Hour


Satisfaction
Violation Pattern
Audit Log Modification
Access Pattern
Data Exfiltration
Privilege Ch
Hour

Appendix K

Institutional Review Board Exemption Letter

NOVA SOUTHEASTERN UNIVERSITY
Institutional Review Board

MEMORANDUM

To: **Karla Clarke**

From: **Ling Wang, Ph.D.,**
Center Representative, Institutional Review Board

Date: **July 20, 2017**

Re: **IRB #: 2017-452; Title, "Novel Alert Visualization: The Development of a Visual Analytics Prototype for Mitigation of Malicious Insider Cyber Threats"**

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review under **45 CFR 46.101(b) (Exempt Category 2)**. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms, they must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE EVENTS/UNANTICIPATED PROBLEMS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and Ling Wang, Ph.D., respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Yair Levy, Ph.D.
Ling Wang, Ph.D.

3301 College Avenue • Fort Lauderdale, Florida 33314-7796
(954) 262-0000 • 800-672-7223, ext. 5369 • Email: *irb@nova.edu* • Web site: www.nova.edu/irb

References

ACI Universal Payments. (2015). Ensures maximum security & compliance. Retrieved May 24, 2016, from http://www.aciworldwide.com/-/media/files/collateral/preventing-money-laundering-and-bank-fraud-in-the-banking-industry-cs-us.pdf

Addae, J., Radenkovic, M., Sun, X., & Towey, D. (2016). An augmented cybersecurity behavioral research model. *IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, 602-603. doi:10.1109/compsac.2016.52

Agrafiotis, I., Nurse, J., Buckley, O., Legg, P., Creese, S., & Goldsmith, M. (2015). Identifying attack patterns for insider threat detection. *Computer Fraud & Security, 7*, 9-17.

Alahmadi, B. A., Legg, P. A., & Nurse, J. R. (2015). Using Internet activity profiling for insider-threat detection. *Proceedings of the 17th International Conference on Enterprise Information Systems*, 1-12. doi:10.5220/0005480407090720

Albanese, M., Pugliese, A., & Subrahmanian, V. (2013). Fast activity detection: Indexing for temporal stochastic automaton-based activity models. *IEEE Transactions on Knowledge and Data Engineering, 25*(2), 360-373. doi:10.1109/TKDE.2011.246

AlMutairi, A., Abdullah, R., AlBukhary, T., & Kar, J. (2015). Security and privacy big data in various applications. *International Journal of Big Data Security Intelligence, 2*(1), 19-24. doi:doi.org/10.14257/ijbdsi.2015.2.1.03

Alvarez, R. (2002). Confessions of an information worker: A critical analysis of information requirements discourse. *Information and Organization, 12*(2), 85–107.

Anderson, C. L., & Agarwal, R. (2010). Practicing safe computing: a multimedia empirical examination of home computer user security behavioral intentions. *MIS Quarterly, 34(*3), 613-643.

Arias-Hernández, R., Dill, J., Fisher, B., & Green, T. (2011). Visual analytics and human-computer interaction. *Interactions, 18*(1), 51-55.

Axelrod, C. W. (2006). Cybersecurity and the critical infrastructure: Looking beyond the perimeter. *Information Systems Control Journal, 3*, 24.

Azaria, A., Richardson, A., Kraus, S., & Subrahmanian, V. (2014). Behavioral analysis of insider threat: A survey and bootstrapped prediction in imbalanced data. *IEEE Transactions on Computational Social Systems, 1*(2), 135-155.

Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management science*, *29*(5), 530-545.

Ball, D. M., & Levy, Y. (2008). Emerging educational technology: Assessing the factors that influence instructors' acceptance in information systems and other classrooms. *Journal of Information Systems Education*, *19*(4), 431.

Ballou, B., Heitger, D. L., & Donnell, L. (2010). Creating effective dashboards: How companies can improve executive decision-making and board oversight. *Strategic Finance*, *91*(9), 27-33.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, *24*(6), 574-594.

Bano, M., Zowghi, D., & da Rimini, F. (2017). User satisfaction and system success: an empirical exploration of user involvement in software development. *Empirical Software Engineering*, *22*(5), 2339-2372.

Big Data Meets Big Data Analytics - SAS. (2012). Retrieved June 28, 2016, from http://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/big-data-meets-big-data-analytics-105777.pdf

Bissell, K. (2013). A strategic approach to cybersecurity. *Financial Executive, 29*(2), 36-42.

Blackwell, D., & Hodges Jr, J. L. (1957). Design for the control of selection bias. *The Annals of mathematical statistics*, 449-460.

Borsci, S., Federici, S., Bacci, S., Gnaldi, M., & Bartolucci, F. (2015). Assessing user satisfaction in the era of user experience: Comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. *International Journal of Human-Computer Interaction*, *31*(8), 484-495.

Boukri, K., & Chaoui, H. (2015). Security analytics in big data infrastructures. *International Journal of Computer Science and Information Security, 13*(5), 91-95.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American educational research journal*, 437-474.

Brdiczka, O., Liu, J., Price, B., Shen, J., Patil, A., Chow, R., ... & Ducheneaut, N. (2012, May). Proactive insider threat detection through graph learning and psychological context. In *Security and Privacy Workshops (SPW), 2012 IEEE Symposium on* (pp. 142-149). IEEE.

Callegati, F., Giallorenzo, S., Melis, A., & Prandini, M. (2017). Insider threats in emerging mobility-as-a-service scenarios.

Cao, Y., Li, Y., Coleman, S., Belatreche, A., & McGinnity, T. M. (2015). Adaptive hidden Markov model with anomaly states for price manipulation detection. *IEEE transactions on neural networks and learning systems, 26*(2), 318-330.

Cappelli, D. M., Moore, A. P., & Trzeciak, R. F. (2012). *The CERT guide to insider threats: How to prevent, detect, and respond to information technology crimes.* Addison-Wesley Professional.

Caputo, D., Maloof, M., & Stephens, G. (2009). Detecting insider theft of trade secrets. *IEEE Security & Privacy Magazine, 7*(6), 14-21. doi: 10.1109/MSP.2009.110

Carcary, M. (2012). IT risk management: A capability maturity model perspective. *Electronic Journal Information Systems Evaluation Volume*, *16*(3).

Casey, E. (2000). *Digital evidence and computer crime.* London: Academic Press.

Casey, T. (2007). *Threat agent library helps identify common security risks.* Retrieved from https://communities.intel.com/thread/49315: https://communities.intel.com/docs/DOC-1151

Casey, T. (2015). *A field guide to insider threat*. tech. rep., Intel.

Casey, T., Koeberl, P., & Vishik, C. (2010, April). Threat agents: a necessary component of threat analysis. In *Proceedings of the sixth annual workshop on cyber security and information intelligence research* (p. 56). ACM.

Chang, L. Y., Zhong, L. Y., & Grabosky, P. N. (2018). Citizen co-production of cyber security: Self-help, vigilantes, and cybercrime. *Regulation & Governance*, *12*(1), 101-114.

Choi, H., Lee, H., & Kim, H. (2009). Fast detection and visualization of network attacks on parallel coordinates. *Computers & Security, 28*(5), 276-288.

Chou, J., Wang, Y., & Ma, K. (2016). Privacy preserving event sequence data visualization using a Sankey diagram-like representation. *SIGGRAPH ASIA 2016 Symposium on Visualization on - SA '16*. doi:10.1145/3002151.3002153

Chouhan, P., & Richhariya, V. (2015). A survey: Analysis of current approaches in anomaly detection. *International Journal of Computer Applications,111*(17), 32-36.

Cole, E. (2015). Insider threats and the need for fast and directed response. *SANS*. Retrieved from http://lp.spectorsoft.com/corp/sans-survey-report

Coursaris, C. K., & Osch, W. V. (2016). A cognitive-affective model of perceived user satisfaction (CAMPUS): The complementary effects and interdependence of usability and aesthetics in IS design. *Information & Management,53*(2), 252-264. doi:10.1016/j.im.2015.10.003

Dalkey N., & Helmer O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science, 9*(3), 458-467.

Dees, P. (2009). Putting the "dash" in the dashboard. *The Journal of the American Society of Military Comptrollers, 54*(1), 8-13.

Denning, D. E., & MacDoran, P. F. (1996). Location-based authentication: Grounding cyberspace for better security. *Computer Fraud & Security*, *1996*(2), 12-16.

DeVita, M. (2005). Medical emergency teams: deciphering clues to crises in hospitals. *Critical Care, 9*(4), 325-326.

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly,18*(4), 453. doi:10.2307/249524

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS quarterly*, 453-461.

Dooley, P. P., Levy, Y., Hackney, R. A., & Parrish, J. L. (2017). Critical value factors in business intelligence systems implementations. In *Analytics and Data Science* (pp. 55-78). Springer, Cham.

Dopping-Hepenstal, L. (1981). Head-up displays. The integrity of flight information. *IEEE Proceedings F Radar and Signal Processing, 128*(7), 440.

Dork, M., Carpendale, S., & Williamson, C. (2011). Visualizing explicit and implicit relations of complex information spaces. *Information Visualization, 11*(1), 5-21. doi:10.1177/1473871611425872

Doskey, S., Mazzuchi, T., & Sarkani, S. (2015). A measure of systems engineering effectiveness in acquisition of complex information systems: A bayesian belief network approach. *IEEE Systems Journal*, *9*(2), 442-450.

Dutta, S., Maeder, A. J., & Basilakis, J. (2013). Using fuzzy logic for decision support in vital signs monitoring. *Joint Proceedings of AIH/CARE* (pp. 29-33).

Ellis, T. J., & Levy, Y. (2009). Towards a guide for novice researchers on research methodology: Review and proposed methods. *Issues in Informing Science and Information Technology, 6*, 323-337.

Ellis, T. J., & Levy, Y. (2010, June). A guide for novice researchers: Design and development research methods. *Proceedings of Informing Science & IT Education Conference, InSITE*, (pp. 108-118).

Elo, S., Kaariainen, M., Kanste, O., Polkki, T., Utriainen, K., & Kyngas, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open,4*(1). doi:10.1177/2158244014522633

Evans, K., & Reeder, F. (2010). *A human capital crisis in cybersecurity: Technical proficiency matters*. CSIS.

Fiebig, A., Halbrügge, M., & Kraus, L. (2016). User experience measurement of a static website compared to a responsive website using attrakdiff mini. *5th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS 2016).* doi:10.21437/pqs.2016-18

Fink, G. A., North, C. L., Endert, A., & Rose, S. (2009, October). Visualizing cyber security: Usable workspaces. *6th International Workshop on Visualization for Cyber Security*, (pp. 45-56).

Foltz, B. (2004). Cyberterrorism, computer crime, and reality. *Information Management & Computer Security, 12*(2), 154-166.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137-144. doi:10.1016/j.ijinfomgt.2014.10.007

Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security, 28*(1), 18-28.

Geer, D. E. (2011). Small is beautiful, big is inevitable. *IEEE Security & Privacy Magazine, 9*(6), 86-87. doi:10.1109/msp.2011.174

Girardin, L., & Brodbeck, D. (1998, December). A visual Approach for monitoring logs. In *LISA* (Vol. 98, pp. 299-308).

Goodall, J. R. (2007). Introduction to visualization for computer security. *VizSEC 2007 Mathematics and Visualization*, 1-17. doi:10.1007/978-3-540-78243-8_1

Gorg, C., Kang, Y., Liu, Z., & Stasko, J. (2013). Visual analytics support for intelligence analysis. *IEEE Computer Society, 46*(7), 30-38. doi:10.1109/MC.2013.76

Gratian, M., Bandi, S., Cukier, M., Dykstra, J., & Ginther, A. (2018). Correlating human traits and cyber security behavior intentions. *Computers & Security*, *73*, 345-358.

Gray, P., & Hovav, A. (2014). Using scenarios to understand the frontiers of IS. *Information Systems Frontiers*, *16*(3), 337-345.

Greitzer, F., & Hohimer, R. (2011). Modeling human behavior to anticipate insider attacks. *Journal of Strategic Security, 4*(2), 25-48. doi:10.5038/1944-0472.4.2.2

Greitzer, F., Moore, A., Cappelli, D., Andrews, D., Carroll, L., & Hull, T. (2008). Combating the insider cyber threat. *IEEE Security & Privacy Magazine, 6*(1), 61-64. doi:10.1109/MSP.2008.8

Hackney, R. A., Dooley, P., Levy, Y., & Parrish, J. (2015). Critical value factors in business intelligence systems implementation success: An empirical analysis of system and information quality. *In Proceedings of the International Conference on Information Systems ICIS2015-SIGDSA* (pp.1-17).

Hadlington, L. (2017). Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours. *Heliyon*, *3*(7), e00346.

Harknett, R. J., & Stever, J. A. (2011). The new policy world of cybersecurity. *Public Administration Review, 71*(3), 455-460. doi:10.1111/j.1540-6210.2011.02366.x

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. *Computers in Human Behavior*, *61*, 463-471.

Harries, A. D., Zachariah, R., Kapur, A., Jahn, A., & Enarson, D. A. (2009). The vital signs of chronic disease management. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 103*(6), 537-540. doi:10.1016/j.trstmh.2008.12.008

Heckman, K., Stech, F., Schmoker, B., & Thomas, R. (2015). Denial and deception in cyber defense. *Computer, 48*(4), 36-44. doi:10.1109/MC.2015.104

Ho, C., Lai, Y., Chen, I., Wang, F., & Tai, W. (2012). Statistical analysis of false positives and false negatives from real traffic with intrusion detection/prevention systems. *IEEE Communications Magazine, 50*(3), 146-154. doi:10.1109/mcom.2012.6163595

Hong, J. C., Tai, K. H., Hwang, M. Y., Kou, Y. C., & Chen, J. S. (2017). Internet cognitive failure relevant to users' satisfaction with content and interface design to reflect continuance intention to use a government e-learning system. *Computers in Human Behavior,66*, 353-362.

Hsu, C. C., & Sandford, B. A. (2007). Minimizing non-response in the Delphi process: how to respond to non-response. *Practical Assessment, Research & Evaluation*, *12*(17), 62-78.

Hueca, A. L., Clarke, K., Levy, Y. (2016). Exploring the motivation behind cybersecurity insider threat and proposed research agenda. *Exploring the motivation behind cybersecurity insider threat and proposed research agenda*, *29*.

Identifiable Information (PII) – NIST. (2010, April). Retrieved November 5, 2016, from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

Imel, S. (2011). Writing a literature review. *Chapter*, *11*, 145.

Inibhunu, C., Langevin, S., Ralph, S., Kronefeld, N., Soh, H., Jamieson, G. A., ... & White, M. (2016). Adapting level of detail in user interfaces for cybersecurity operations. *In Resilience Week (RWS)*, (pp. 13-16).

Jackson, G. M. (2012). *Predicting malicious behavior: Tools and techniques for ensuring global security*. John Wiley & Sons.

Jonas, J. (2006). Threat and fraud intelligence, Las Vegas style. *IEEE Security and Privacy Magazine, 4*(6), 28-34.

Julisch, K. (2003). Clustering intrusion detection alarms to support root cause analysis. *ACM transactions on information and system security, 6*(4), 443-471.

Jyothsna, V., Prasad, V. R., & Prasad, K. M. (2011). A review of anomaly based intrusion detection systems. *International Journal of Computer Applications, 28*(7), 26-35.

Kandel, S., Paepcke, A., Hellerstein, J. M., & Heer, J. (2012). Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics, 18*(12), 2917-2926.

Kang, Y., Gorg, C., & Stasko, J. (2011). How can visual analytics assist investigative analysis? Design implications from an evaluation. *IEEE Transactions on Visualization and Computer Graphics, 17*(5), 570-583. doi:10.1109/tvcg.2010.84

Keeney, R. L. (1999). The value of Internet commerce to the customer. *Management Science,45*(4), 533-542. doi:10.1287/mnsc.45.4.533

Keim, D. (2000). Designing pixel-oriented visualization techniques: Theory and applications. *IEEE Transactions on Visualization and Computer Graphics, 6*(1), 59-78. doi:10.1109/2945.841121

Kemmerer, R., & Vigna, G. (2002). Intrusion detection: A brief history and overview. *Computer, 35*(4), 27-30.

Kim, K. (1989). User satisfaction: A synthesis of three different perspectives. *Journal of Information Systems*, *4*(1), 1-12.

King, Z., Henshel, D., Flora, L., Cains, M. G., Hoffman, B., & Sample, C. (2018). Characterizing and Measuring Maliciousness for Cybersecurity Risk Assessment. *Frontiers in Psychology*, *9*, 39.

Kott, A., Swami, A., & Mcdaniel, P. (2014). Security outlook: Six cyber game changers for the next 15 years. *Computer, 47*(12), 104-106. doi:10.1109/mc.2014.366

Kumar, K. A., & NandaMohan, D. V. (2008). Novel anomaly intrusion detection using neuro-fuzzy inference system. *International Journal of Computer Science and Network Security, 8*(8), 6-11.

Kumarmandal, K., & Chatterjee, D. (2015). Insider threat mitigation in cloud computing. *International Journal of Computer Applications, 120*(20), 7-11.

Kurucay, M., & Inan, F. A. (2017). Examining the effects of learner-learner interactions on satisfaction and learning in an online undergraduate course. *Computers & Education*, 115, 20-37.

Lazarevic, A., Kumar, V., & Srivastava, J. (2005). Intrusion detection: A survey. In *Managing Cyber Threats* (pp. 19-78). Springer US.

Lee, S. M., Kim, Y. R., & Lee, J. (1995). An empirical study of the relationships among end-user information systems acceptance, training, and effectiveness. *Journal of management information systems*, *12*(2), 189-202.

Legg, P., Buckley, O., Goldsmith, M., & Creese, S. (2015). Automated insider threat detection system using user and role-based profile assessment. *IEEE Systems Journal, PP*(99), 1-10.

Legg, P., Moffat, N., Nurse, J., Happa, J., Agrafiotis, I., Goldsmith, M., & Creese, S. (2013). Towards a conceptual model and reasoning structure for insider threat detection. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 4*(4), 20-37.

Leventhal, B. (2010). An introduction to data mining and other techniques for advanced analytics. *Journal of Direct, Data and Digital Marketing Practice, 12*(2), 137-153.

Levy, Y. (2006). *Assessing the value of e-learning systems*. IGI Global.

Levy, Y. (2008). An empirical development of critical value factors (CVF) of online learning activities: An application of activity theory and cognitive value theory. *Computers & Education*, *51*(4), 1664-1675.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: International Journal of an Emerging Transdiscipline, 9*(1), 181-212.

Levy, Y., & Ellis, T. J. (2011). A guide for novice researchers on experimental and quasi-experimental studies in information systems research. *Interdisciplinary Journal of Information, Knowledge, and Management, 6*, 151-161.

Levy, Y., & Ramim, M. M. (2004). Financing expensive technologies in an era of decreased funding: Think Big... Start Small... Build Fast... In C. Howard, K. Schuenk, & R. Discenza (Eds.), *Distance Learning and University Effectiveness: Changing Educational Paradigms for Online Learning* (pp. 278-301). Hershey, PA : Idea-Group Publishing.

Levy, Y., & Ramim, M. M. (2012). A study of online exams procrastination using data analytics techniques. *Interdisciplinary Journal of E-Learning and Learning Objects*, *8*(1), 97-113.

Levy, Y., Murphy, K. E., & Zanakis, S. H. (2010). A value-satisfaction taxonomy of IS effectiveness (VSTISE): A Case Study of User Satisfaction with. *Information Systems and New Applications in the Service Sector: Models and Methods: Models and Methods*, 90.

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. *Human Centered Design Lecture Notes in Computer Science*, 94-103. doi:10.1007/978-3-642-02806-9_12

Malec, P., Piwowar, A., Kozakiewicz, A., & Lasota, K. (2015). Detecting security violations based on multilayered event log processing. *Journal of Telecommunications and Information Technology*, (4), 30.

Mason, C. H., & Perreault Jr, W. D. (1991). Collinearity, power, and interpretation of multiple regression analysis. *Journal of marketing research*, 268-280.

McKenna, S., Mazur, D., Agutter, J., & Meyer, M. (2014). Design activity framework for visualization design. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 2191-2200.

McKenna, S., Staheli, D., & Meyer, M. (2015). Unlocking user-centered design methods for building cyber security visualizations. *IEEE Symposium on Visualization for Cybersecurity (VizSec)*, (pp. 1-8).

Mell, P., Bergeron, T., & Henning, D. (2005, November). Creating a patch and vulnerability management program. Retrieved June 28, 2016, from http://csrc.nist.gov/publications/nistpubs/800-40-Ver2/SP800-40v2.pdf

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods*. Glendale.

Miller, H. G., & Murphy, R. H. (2009). Secure cyberspace: Answering the call for intelligent action. *IT professional, 11*(3), 60-63.

Mok, W. Q., Wang, W., & Liaw, S. Y. (2015). Vital signs monitoring to detect patient deterioration: An integrative literature review. *International Journal of Nursing Practice, 21*, 91-98. doi:10.1111/ijn.12329

National Institute of Standards and Technology (NIST). (2013). Glossary of key information security terms (NISTIR 7298, Revision 2). Retrieved from http://nvlpubs.nist.gov/nistpubs/ir/2013/NIST.IR.7298r2.pdf

NIST: 800-122 Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). (2010, August). Retrieved November 5, 2016, from http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

NIST: 800-61 Computer Security Incident Handling Guide. (2012, August). Retrieved November 5, 2016, from http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf

Noel, S. (2015). Interactive visualization and text mining for the CAPEC cyber attack catalog. In *Proceedings of the ACM Intelligent User Interfaces Workshop on Visual Text Analytics*.

Nostro, N., Ceccarelli, A., Bondavalli, A., & Brancati, F. (2013). *A methodology and supporting techniques for the quantitative assessment of insider threats. Proceedings of the 2nd International Workshop on Dependability Issues in Cloud Computing - DISCCO 13*. doi:10.1145/2506155.2506158

Okesola, J. O., Ogunseye, O. S., & Folorunso, O. (2010). An efficient multi-expert knowledge capture technique. *International Journal of Computer Applications, 8*(10). doi:http://dx.doi.org.ezproxylocal.library.nova.edu/10.5120/1245-1611

Ozcan, M. B., & Morrey, I. (1995). A visual requirements validation environment for the reverse engineering of formal specifications from rapid prototypes. *ACM SIGSOFT Software Engineering Notes,20*(5), 83-87. doi:10.1145/217030.217047

Patan, K. (2015). Neural network-based model predictive control: Fault tolerance and stability. *IEEE Transactions on Control Systems Technology, 23*(3), 1147-1155.

Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks, 51*(12), 3448-3470.

Petter, S., Delone, W., & Mclean, E. R. (2013). Information systems success: The quest for the independent variables. *Journal of Management Information Systems,29*(4), 7-62. doi:10.2753/mis0742-1222290401

Pfleeger, S., & Stolfo, S. (2009). Addressing the insider threat. *IEEE Security & Privacy Magazine, 7*(6), 10-13.

Pfleeger, S., Predd, J., Hunker, J., & Bulford, C. (2010). Insiders behaving badly: Addressing bad actors and their actions. *IEEE Transactions on Information Forensics and Security, 5*(1), 169-179.

Pinsonneault, A., & Kraemer, K. (1993). Survey research methodology in management information systems: An assessment. *Journal of Management Information Systems,10*(2), 75-105. doi:10.1080/07421222.1993.11518001

Predd, J., Pfleeger, S., Hunker, J., & Bulford, C. (2008). Insiders behaving badly. *IEEE Security & Privacy, 6*(4), 66-70. doi:10.1109/MSP.2008.87

Price WaterHouse Coopers (2013). Key findings from the 2103 U.S. State of cybercrime survey

Price WaterHouse Coopers (2014). Managing cyber risks in an interconnected world: Key findings from the global state of information security

Qayyum, A., Islam, M., & Jamil, M. (2005). Taxonomy of statistical based anomaly detection techniques for intrusion detection. *Proceedings of the IEEE Symposium on Emerging Technologies, 2005*. 270-276. doi:10.1109/icet.2005.1558893

Ramdhani, A., Ramdhani, M. A., & Amin, A. S. (2014). Writing a literature review research paper: A step-by-step approach. *International Journal of Basic and Applied Science*, *3*(1), 47-56.

Ramim, M. M., & Lichvar, B. T. (2014). Eliciting expert panel perspective on effective collaboration in system development projects. *Online Journal of Applied Knowledge Management*, *2*(1), 122-136.

Randazzo, M. R., Keeney, M., Kowalski, E., Cappelli, D., & Moore, A. (2005). Insider threat study: Illicit cyber activity in the banking and finance sector (No. CMU/SEI-2004-TR-021). Carnegie-Mellon University Pittsburg PA Software Engineering Institution

Rantala, R. (2008). Cybercrime against businesses, 2005, special report NCJ221943, U.S. Bureau of Justice Statistics. Retrieved June 30, 2016, from http://www.bjs.gov/content/pub/pdf/cb05.pdf

Roberts, J., Ritsos, P., Badam, S., Brodbeck, D., Kennedy, J., & Elmqvist, N. (2014). Visualization beyond the desktop--the next big thing. *IEEE Computer Graphics and Applications, 34*(6), 26-34. doi:10.1109/MCG.2014.82

Rogers, A. E., Dean, G. E., Hwang, W., & Scott, L. D. (2008). Role of registered nurses in error prevention, discovery and correction. *Quality and Safety in Health Care, 17*(2), 117-121. doi:10.1136/qshc.2007.022699

Rohrer, R. M., & Swing, E. (1997). Web-based information visualization. *IEEE Computer Graphics and Applications, 17*(4), 52-59.

Rosemann, M., & vom Brocke, J. (2015). The six core elements of business process management. In *Handbook on business process management 1* (pp. 105-122). Springer Berlin Heidelberg.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, *15*(4), 353-375.

Sanders, W. H. (2017, September). Data-driven model-based detection of malicious insiders via physical access logs. In *Quantitative Evaluation of Systems: 14th International Conference, QEST 2017, Berlin, Germany, September 5-7, 2017, Proceedings* (Vol. 10503, p. 275). Springer.

Santos, E., Nguyen, H., Yu, F., Kim, K., Li, D., Wilkinson, J., . . . Clark, B. (2012). Intelligence analyses and the insider threat. *IEEE Transactions on Systems Management and Cybernetics, 42*(2), 331-347.

Sedera, D., Lokuge, S., Grover, V., Sarker, S., & Sarker, S. (2016). Innovating with enterprise systems and digital platforms: A contingent resource-based theory view. *Information & Management*, *53*(3), 366-379.

Sekaran, U., & Bougie, J. R. (2009). Research methods for business: A skill-building approach (6th ed.). Chichester: John Wiley & Sons.

Shami, N. S., Muller, M., Pal, A., Masli, M., & Geyer, W. (2015, April). Inferring employee engagement from social media. *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3999-4008). ACM.

Sherman, D. C. (2013). *US cybersecurity defense assessment*. Army war college Carlisle Barracks PA.

Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on* (pp. 336-343). IEEE.

Shneiderman, B., & Plaisant, C. (2015). Sharpening analytic focus to cope with big data volume and variety. *IEEE Computer Graphics and Applications, 35*(3), 10-14. doi:10.1109/mcg.2015.64

Shneiderman, B., Plaisant, C., Cohen, M., & Jacobs, S. (2010). *Designing the user interface: Strategies for effective human-computer interaction.* Boston: Addison-Wesley.

Sigaard, K. T., & Skov, M. (2015). Applying an expectancy-value model to study motivators for work-task based information seeking. *Journal of Documentation*, *71*(4), 709-732.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *Journal of Information Technology Education*, *6*(1), 1-21.

Solms, R. V., & Niekerk, J. V. (2013). From information security to cyber security. *Computers & Security, 38*, 97-102. doi:10.1016/j.cose.2013.04.004

Spathoulas, G. P., & Katsikas, S. K. (2010). Reducing false positives in intrusion detection systems. *Computers & Security, 29*(1), 35-44. doi:10.1016/j.cose.2009.07.008

St Louis, K. O., Lubker, B. B., Yaruss, J. S., & Aliveto, E. F. (2009). Development of a prototype questionnaire to survey public attitudes toward stuttering: Reliability of the second prototype. *Contemporary Issues in Communication Sciences and Disorders*, *36*, 101-107.

Staheli, D., Mancuso, V. F., Leahy, M. J., & Kalke, M. M. (2016). Cloudbreak: Answering the challenges of cyber command and control. *Lincoln Laboratory Journal, 22*(1).

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

Straub, D. W. (1989). Validating instruments in MIS research. *MIS Quarterly,13*(2), 147. doi:10.2307/248922

Sun, L., Srivastava, R. P., & Mock, T. J. (2006). An information systems security risk assessment model under the Dempster-Shafer theory of belief functions. *Journal of Management Information Systems*, *22*(4), 109-142.

Ten, C. W., Manimaran, G., & Liu, C. C. (2010). Cybersecurity for critical infrastructures: attack and defense modeling. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 40*(4), 853-865.

The White House. (2009, March). Retrieved November 5, 2016, from https://www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf

Thomson, K., & Solms, R. V. (2005). Information security obedience: A definition. *Computers & Security, 24*(1), 69-75. doi:10.1016/j.cose.2004.10.005

Toecker, M. (2014). Generation cybersecurity: What you should know, and be doing about it. *Power, 158*(2), 40-45.

Turker, D., Turker, D., Ozmen, Y. S., & Ozmen, Y. S. (2017). Linking values and ideologies: a scale of managerial social responsibility values. *Journal of Global Responsibility*.

Van Teijlingen, E. R., & Hundley, V. (2001). The importance of pilot studies.

Vera-Baquero, A., Colomo Palacios, R., Stantchev, V., & Molloy, O. (2015). Leveraging big-data for business process analytics. *The Learning Organization*, *22*(4), 215-228.

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: effects of activation and self-centrality of values on choices and behavior. *Journal of personality and social psychology*, *82*(3), 434.

Victor, G. J., Rao, D. M., & Venkaiah, D. V. (2010). Intrusion detection systems - analysis and containment of false positives alerts. *International Journal of Computer Applications, 5*(8), 27-33. doi:10.5120/931-1308

Walton, S., Maguire, E., & Chen, M. (2015, October). A visual analytics loop for supporting model development. *IEEE Symposium on Visualization for Cybersecurity (VizSec)*, (pp. 1-8).

Wang, J., & Paschalidis, I. C. (2015). Statistical traffic anomaly detection in time-varying communication networks. *IEEE Transactions on Control of Network Systems, 2*(2), 100-111.

Wang, L., & Jones, R. (2017). Big data analytics for network intrusion detection: A survey. *International Journal of Networks and Communications*, *7*(1), 24-31.

Warkentin, M., & Willison, R. (2009). Behavioral and policy issues in information systems security: the insider threat. *European Journal of Information Systems, 18*(2), 101.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, 26(2), xiii-xxiii.

Westerlund, M., Craigen, D., Bailetti, T., & Agwae, U. (2018). A three-vector approach to blind spots in cybersecurity. In *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 1684-1693). IGI Global.

Xu, J., You, J., & Liu, F. (2005, March). A fuzzy rules based approach for performance anomaly detection. In Proceedings. *IEEE Networking, Sensing and Control*, (pp. 44-48).

Ye, N., Emran, S. M., Chen, Q., & Vilbert, S. (2002). Multivariate statistical analysis of audit trails for host-based intrusion detection. *IEEE Transactions on Computers, 51*(7), 810-820.

Yoo, S., Jo, J., Kim, B., & Seo, J. (2018). LongLine: Visual analytics system for large-scale audit logs. *Visual Informatics*, *2*(1), 82-97.

Yu, Y. (2012). A survey of anomaly intrusion detection techniques. *Journal of Computing Sciences in Colleges, 28*(1), 9-17.

Zaknich, A. (1998). Introduction to the modified probabilistic neural network for general signal processing applications. *IEEE Transactions on Signal Processing, 46*(7), 1980-1990.

Zhang, J., & Zulkernine, M. (2006, June). Anomaly based network intrusion detection with unsupervised outlier detection. *IEEE International Conference on Communications, 5*, 2388-2393.