**Improved Detection and Response Via Optimized Alerts: Usability Study**

Russell McRee, PhD

Supplement: Literature Review

The focal points for this review of the current academic literature will be as follows: organizations risk data breach, loss of valuable human resources, reputation, and revenue due to excessive security alert volume and a lack of fidelity in security alert output. Where security alerts generated via data science and machine learning methods can be central to this problem, the opportunities for improved security alert output from these methods are key. Most importantly, the security analyst’s perception of the usability of the resulting security alert output generated via these methods is pivotal. This chapter will be composed of sections to support this assessment of current literature. First, the process and criteria used to search the related scholarly literature will be described. Second, an overview of the theoretical orientation for this study will follow. Finally, there will be a review of the existing academic literature itself, composed of the following focus areas:

* Security Information and Event Management (SIEM)
* Security Analyst Challenges: Scale, Volume, Noise, and Workload
* Security Data Analysis and Processing
* Machine Learning (ML) and Data Science (DS)
* Security Data Visualization

The review will be followed by key findings identified during the literature review, a critique of previous research methods, and finally a summary and a transition to Chapter 3.

# Methods of Searching

The search for relevant scholarly literature will include a variety of sources and methods. The effective use of keywords, Boolean search operators, and filtering will be key to success. Searches often begin with online search engines to achieve a broader, then are refined with university library sources. Search phrases will include the likes of *information security analyst data fatigue filetype:pdf* and *high volume of security alert and event data*. Keyword term searches will include *security alerts*, *security* *alert fatigue*, *security data science*, *machine learning*, *security data visualization*, and *situation awareness*. Boolean search operations included the likes of *(security information and event management) AND (security operations center)*.

Additional benefit will be derived from searching references from selected scholarly literature. Often researchers conducting related research will draw from topic-specific references.

A secondary literature search will target the Technology Acceptance Model (TAM) and its related use as a viable survey instrument for this study. TAM is well established to assess perceived usefulness and perceived ease of use and has been adapted and evolved over time. As such, a review of literature specific to TAM refinement, including questionnaire item order, will prove useful as a way of establishing an optimized survey. As an example, numeric response options arranged with magnitude of agreement increasing from left to right were declared the best choice for future research by Lewis (2019) given similarity to other popular measures of perceived usability such as the System Usability Scale (SUS) and Usability Metric for User Experience (UMUX-LITE).

# Theoretical Orientation for the Study

The theoretical foundation of this study will be the Technology Acceptance Model (TAM) as conceived by Davis (1989). The pillars of TAM are perceived usefulness and perceived ease of use. Davis (1989) stated that perceived usefulness is defined as the degree to which people believe that use of a particular system enhances job performance, while perceived ease of use, in contrast, refers to the degree to which it is believed that use of a particular system would be free of effort. TAM has been applied across a wide spectrum, from analysis of students’ behavioral intention to use e-learning (Park, 2009) and evaluation of consumer perceptions of mobile commerce (Wu & Wang, 2005). Additional applications have included the perceived usefulness and ease of use of electronic supermarkets (Henderson & Divett, 2003) and a comparison of differences in user acceptance models for productivity-oriented (utilitarian) and pleasure-oriented (hedonic) information systems (Van der Heijden, 2004).

The theory of reasoned action (TRA), a well-established model used to predict and explain human behavior in a variety of venues, was the precursor to Davis’ TAM (Wu et al., 2007). TAM has been tested and extended by numerous empirical research studies (Davis, 1989; Henderson & Divett, 2003; Igbaria, Zinatelli, Cragg, & Cavaye, 1997; Legris, Ingham, & Collerette, 2003; Venkatesh & Davis, 2000). Davis (1989) states that the TAM model consists of perceived ease of use (PEU), perceived usefulness (PU), intention to use (IU), and attitude toward using (AU). While PU and PEU are considered primary indicators for system use, prior research has indicated that attitude towards the technology is not as significant a contributor (Wu et al., 2007). TAM has been proven for its validity and ability to satisfactorily explain end user system usage.

User satisfaction in the era of user experience is central to the modern analysis of human-computer interaction (HCI; Borsci et al., 2015). It is generally agreed that the user’s interactive experience is affected by the perceived usability and aesthetics of an interface (Hassenzahl, 2005; Hassenzahl & Tractinsky, 2006). In the context of HCI and user experience, Hassenzahl (2005) asserted that products should include features such as content, presentational style, functionality, and interactional style. These should be selected by designers to summarize certain attributes: novel, interesting, useful, and predictable. For the purposes of this study, the products to be considered in this light are security alert outputs, both text-based and visualized. Hassenzahl and Tractinsky (2006) again indicated that the perceived affective quality of a system is an antecedent of its perceived usability, usefulness, and the intention to use. These finding are in keeping with previous research where usefulness and enjoyment were found to directly influence the effects on usage intentions as they relate to perceived output quality and perceived ease of use (Davis et al., 1992). For the security analyst faced with security data fatigue, the requirement to increase security event detection accuracy, perceived output quality and perceived ease of use are key. Where security analysts benefit from enhanced experiences with security alert data as generated via data science and machine learning models, the quality of the output will be paramount to that experience.

# Review of the Literature

This literature review will follow a logical sequence intended to illuminate the lifecycle of the information security analyst charged with analyzing and responding to security events and alerts at scale. To thoroughly understand the related challenges, and ultimately, the business problem these analysts and their organizations face, numerous perspectives will be explored. The inception of most security event and alert processes begins with security incident and event management, commonly both a platform and a practice, often referred to with the industry acronyms SIEM or SEIM. These systems often generate significant challenges for analysts, represented in issues of data scale, alert volumes, signal to noise ratio, and excess workloads. The methods by which these issues and challenges are addressed are best described as security data analysis and processing. These methods can be greatly accentuated and optimized with the use of machine learning and data science. Finally, the output of these accentuated and optimized methods (models) can be represented via the practice of security data visualization.

### Security Information and Event Management

To appreciate the challenges organizations and their security analysts face the review must begin with the practice central to most security analysis and incident response processes, namely security information and event management (SIEM). Where organizations risk data breach, loss of valuable human resources, reputation, and revenue due to excessive security alert volume and a lack of fidelity in security event data, the challenges often arise in the context of the SIEM program and process. Security managers utilize SIEMs to support early attack detection, investigation, and response to identify and prevent security breaches from security event data in real time (Kavanagh, Bussa, & Sadowski, 2020). SIEMs aggregate, correlate, identify, detect, investigate, and report on security data to ensure that requirements relating to incident response, forensics and regulatory compliance are met (Kavanagh, et al., 2020). Organizations and enterprises are targeted by adversaries, criminals, and Advanced Persistent Threats (APTs), requiring the deployment of intrusion detection systems (IDS) and SIEMs to monitor the enterprise-wide activity and generate security alerts specific to suspicious activity (Hassan, et al., 2019). Security analysts triage and process these alerts to identify true positive attack signals. These IDS and SIEM systems are prone to a high rate of false positive alerts (Hassan, et al., 2019). SIEM deployments often fail to achieve their full potential due to a lack of integration as part of the larger information technology environments in which they operate, leading to an inability to acquire data necessary to operationalize pre-defined use cases specific to SIEMs for threat detection (Slatman, 2016). Security Operations Centers (SOC) also struggle to meet the technical and operational skill requirements to operate SIEM solutions to the greatest effect (Slatman, 2016).

To understand SIEM use is to understand the lifecycle and workloads of a SOC. SOCs are often organized by tiers, where Tier 1 security analysts represent the front line of defense utilizing the SIEM to monitor their enterprises for attempts to breach security (Sundaramurthy, et al., 2015). Tier 2 security analysts serve in more of a managerial, mentoring, and escalation capacity for the often more junior Tier 1 security analysts. Tier 2’s primary role is to ensure operational visibility for organization leadership metrics, key performance indicators, and reporting (Sundaramurthy, et al., 2015). SOCs are often utilized to provide managed security services (MSS) to provide network security monitoring (NSM) for customers. The SIEM lays at the heart of this process to enable the collection, analysis, and escalation of indicators and warnings to detect and respond to intrusions. SOC Tier 1 analysts correlate alerts to confirm success of attempted exploits. If confirmed as a security incident as confirmed by true positive alerts, Tier 2 analysts commonly escalate findings to decision makers for reaction and response (Khalili, et al., 2019).

SIEM data can be further extended to provide input to additional processes including security assessments and the selection of defensive countermeasures utilizing quantitative metrics (Kotenko & Doynikova, 2017). Data from the SIEM can also be utilized to generate visualizations and attack graphs, as explored later in this chapter under Security Data Visualization.

### Security Analyst Challenges: Scale, Volume, Noise, and Workload

Security analysts often work under high stress (Sundaramurthy et al., 2015). Amongst many challenges organizations face are constantly evolving attacks, where adversaries change their behaviors, and limited investigation time and budgets, where relying on security analysts can be costly and time consuming (Veeramachaneni, et al., 2016). Significant quantities of network traffic logs generated by people and devices are difficult to analyze given the scale; where security analysts seek to detect anomalies in such traffic for potential faults and security breaches, they face difficulty (Liao, et al., 2018).

**Scale.**

Scale can be used to describe two issues security analysts face, both the scale of attacks and breaches, as well as the scale of the networks and computing infrastructure they are challenged with protecting. Large-scale data breaches in the private and public sectors are illuminating the need to enhance cyber defenses (Vieane, et al., 2016), while web-scale platforms generate millions of log entries per day: 3.6 billion log lines over a period of 3 months, as an example (Veeramachaneni, et al., 2016). To monitor web-scale systems, security platforms must be capable of analyzing 10+ million entities daily while also being able to update and retrieve behavioral signatures for these entities on demand and in real time - as many as 50,000 at once (Veeramachaneni, et al., 2016). Meanwhile, large-scale attacks such as distributed denial of service (DDoS) attacks stress both defensive capabilities as well as the requirement for reliable intrusion detection systems (IDS) to ensure low false positives (Moustafa, et al., 2019). Detective systems deployed at the edges of large-scale networks must adapt to global distribution (time differences, varied network speeds), high-mobility, and low-latency processing challenges (Moustafa, et al., 2019). Security analysts operating on behalf of SOCs deal with massive scale and complexity on behalf of their customer’s IT environments, facing significant diversity due to increased mobility, larger scale IT usage, the prevalence of cloud infrastructure and difficulties associated with the bring-your-own-device (BYOD) phenomenon (Slatman, 2016). As a result, these SOCs have less visibility and control over the networks they are responsible for.

**Volume.**

To address issues with data volume, systems must be able to process high-volume, high-velocity data at challenging scale to overcome limited analyst bandwidth (Veeramachaneni et al., 2016). The sheer speed and large size of current network environments leads to network data with the characteristics of big data, defined in terms of:

* Volume: amount of data
* Velocity: speed of data processing
* Variety: complexity, diversity, and dimension of data (Moustafa, et al., 2019)

As organizations experience rapid growth of dynamic network size, volume, dimension, and complexity, data mining and standard analysis methods no longer suffice (Liao, et al., 2018). Security analysts now face a volume of global Internet traffic that as of 2020 is 95 times higher than the volume in 2005 (Cisco, 2017). As a result false positive alerts increase and are themselves a volume problem (Rajivan, 2011). Additionally, malicious attacks are constantly evolving and occur in very large volumes requiring scalable solutions (Vinayakumar et al., 2019). Human cognitive capabilities become challenged to accurately detect cyber-attacks, considering the high volume of security alerts to be processed and the high probability of false alerts (Ben-Asher & Gonzalez, 2015). To address the issues associated with volume, approaches such as anomaly detection can aid security analysts in responding to the high volume of security alerts, additionally helpful them identify events and changes in the environment corresponding to security incidents and malware outbreaks (Pierazzi, et al., 2016). In one 2017 survey, respondents stated that the volume and severity of malware infections increased over the prior year by as much as 36% (Ponemon, 2017). In short, security analysts face the difficult problems of volume and complexity, leading to decreased efficacy specific to situation awareness (Giacobe, 2013)

**Noise.**

Another challenge security analysts face is one of noise in the security alert data. It is common that a certain amount of malicious activity is not systematically reported due to inconclusive incident response or a lack of initial detection (Veeramachaneni et al., 2016). This introducing noise into the data, given that unreported attacks will be considered legitimate (Veeramachaneni et al., 2016). For security analysts to be able to analyze and screen security alert noise and focus on true threats, automation becomes an inherent requirement (Cisco, 2017). As an example, the negative effects on security detection performance caused by system noise can be reduced by processing relevant data with multivariate statistical and Kalman filtering methods (Wu & Wang, 2018). Noise is a significant issue when more than half of security alerts are false positives, and likely even higher in the absence of alert correlation and deduplication where more than one-third of alerts may be redundant (FireEye, 2015). Signal detection theory (SDT) then comes into play, where it becomes critical for the process of security alert analysis to separate meaningful information from noise (Lynn & Barrett, 2014).

**Workload.**

Security analysts and SOCs must additionally increase workload as a function of the scale, volume, and noise that they face daily. Increased workload driven by high event counts is an ideal incentive for automation (Sundaramurthy et al., 2015). Vieane et al. (2016) considered display information density for information security analysts to be several orders of magnitude above that seen in the previously explored medical monitoring and air traffic control domains. Grier (2015), in an analysis of cognitive workload, found that air traffic control, command and control, and medical tasks were considered the most demanding cognitive domains. As such, the requirement to ensure that security analyst workload does not exceed their information processing capacity is essential (Vieane et al., 2016). The cognitive workload for security analysts was impacted by the length of attention required for task completion (Sawyer et al., 2014). In a study specific to network defense, Greenlee et al. (2016) noted that cognitive workload and stress were factors in Tier 1 SOC tasks (initial network intrusion triage) and Tier 2 SOC tasks (escalation analysis). Paul and Dykstra (2017) found that, in the environment specific to the National Security Agency, increased awareness of the factors affecting operator fatigue, frustration, and cognitive workload have driven changes in organization policies and practices. Paul & Dykstra (2017) further provided a baseline against which to measure and compare the effectiveness of interventions devised to mitigate negative effects from the above-mentioned critical human factors.

### Security Data Analysis and Processing

Ben-Asher and Gonzalez (2015) assessed the effects of cyber-security knowledge on attack detection with attention to cyber-security, dynamic decision-making, and intrusion-detection systems. Ben-Asher and Gonzalez (2015) developed a simplified Intrusion Detection System (IDS) to assess individuals with varied levels of knowledge specific to detection of malicious security events, as well as their ability to properly declare an attack based on their analysis of a sequence of network events. Their analysis revealed that most participants in the expert group (80%) worked in network operation and information security for a year or more, and 35% of them had more than 10 years of related on-the-job experience. On the other hand, 93% of novice group participants stated that they have no experience in network operation and information security (Ben-Asher & Gonzalez, 2015). Ben-Asher and Gonzalez (2015) concluded that expertise and practical knowledge play an important role in triage analysis, as well as tasks such as classifying a network event as a threat (or not) and identifying the connections between small decisions and overall attack decisions based on network event sequences (Ben-Asher & Gonzalez, 2015). These findings relate to cognitive processes and have implications for improving cyber security. Importantly, and related to this research, Ben-Asher and Gonzalez (2015) determined that a high volume of intrusion alerts to be processed, coupled with the high probability of false alerts, challenges human cognitive capabilities in accurately detecting an attack (Ben-Asher & Gonzalez, 2015). This research is further relevant in that, in scenarios where analysts are junior and inexperienced, in addition to being overwhelmed with data, tools such as machine learning, automation and visualization should increase their efficacy.

Processing security-related big data is often difficult, and conventional tools and methods are no longer viable or scalable (Vinayakumar et al. 2019). Scalable frameworks present opportunities to address these big data challenges using cloud-scale technologies, hardware, and architecture to expedite data processing. Vinayakumar et al. (2019) suggested a specific framework utilizing distributed computing and distributed machine learning models for effective identification of intrusions and attacks, at both the network-level and host-level. The Scale-Hybrid-IDS-AlertNet (SHIA) framework offers a scalable design as a distributed monitoring and reporting system (Vinayakumar, et al. 2019). SHIA uses characteristics of distributed computing and distributed machine learning algorithms to enhance analysis and processing performance while deploying the hybrid system to different network locations (Vinayakumar et al. 2019).

Dutt et al. (2013) analyzed cyber situation awareness (SA) by modeling detection of cyber-attacks with attention to cyber situation awareness, instance-based learning theory, and adversarial behavior on behalf of defenders. Dutt et al. (2013) utilized different model types to represent a security analyst facing different adversarial behaviors. They did so to “determine the effects of an adversary’s behavior on the defender’s accurate and timely detection of network threats” (Dutt, et al, 2013, p. 2). Their results found that, where an impatient strategy applied, risk-averse models with threat-prone experiences show improved detection compared with risk-seeking models with nonthreat-prone experiences. They also discovered that this, however, is not true for a patient strategy. Dutt et al. (2013) concluded that a defender’s prior threat experiences and his or her tolerance to threats are likely to predict detection accuracy. They also stated that considering the nature of adversarial behavior is also important. Dutt, et al. (2013) recommended that, to improve a defender’s overall threat detection, decision-support tools that consider the role of a defender’s experience and tolerance to threats along with the nature of adversarial behavior are likely to be helpful. A theme related to this research surfaces: decision-making models better support security analysts considering adversarial behavior, thus supporting the argument for improving the security analyst experience with machine learning and security data visualization.

### Machine Learning and Data Science

Hassan et al. (2019) referenced network diffusion algorithms and aggregate anomaly scores. With attention to alert fatigue, threat detection software, and intrusion detection, Hassan et al. (2019) presented NODOZE to combat threat alert fatigue using contextual and historical information of generated threat alerts. Hassan et al. (2019) generated a graph of a security alert event, assigning an anomaly score to each edge in the graph based on the event frequency in the enterprise. Hassan et al. (2019) propagated scores along the neighboring edges of the graph using a network diffusion algorithm, then generated an aggregate anomaly score for triaging use (Hassan et al., 2019). According to the Hassan et al. (2019), NODOZE consistently ranked the true alerts higher than the false alerts based on aggregate anomaly scores. When Hassan et al. (2019) introduced a cutoff threshold for anomaly scores, NODOZE decreased the volume of false alarms by 84%, saving security analysts’ more than 90 hours of investigation time per week. Hassan et al. (2019) found that NODOZE generates alert dependency graphs two orders of magnitude smaller than those generated by traditional tools, while not sacrificing vital information needed for investigation. Hassan et al. (2019) further asserted that NODOZE runtime overhead is low enough to be practical and can be deployed with any threat detection software. Hassan et al. (2019) employed sound logic to decreasing alert volume and improving analyst experience.

Moustafa et al. (2019) explored network anomaly detection systems referencing machine learning principles to describe various Decision Engine (DE) approaches, including new ensemble learning and deep learning approaches. Moustafa et al.’s (2019) research results indicated that, for learning and validating ML mechanisms on new datasets, combination and statistical techniques can effectively detect existing and zero-day attacks while knowledge, classification, and clustering can efficiently detect known attacks. To reach these conclusions, Moustafa et al. (2019) utilized data pre-processing methods for creating, generating, reducing, converting, and normalizing features to pass filtered information to a decision engine, which distinguishes between anomalous and legitimate observations. Moustafa et al. (2019) applied these with classification, clustering, knowledge, combination, and statistics to demonstrate their merits and demerits in terms of building an effective network anomaly detection system. Moustafa, et al. (2019) found that real-time detection is challenging. The features created for network traffic are often noisy or irrelevant, and the weight (cost to perform/calculate) of detection methods must be carefully adopted, with respect to the above problems (Moustafa, et al., 2019). “These reasons increase the processing time and false alarm rate if not properly addressed. Therefore, feature reduction and lightweight decision approaches should be developed” (Moustafa, et al, 2019).

Pierazzi et al. (2016) researched exploratory security analytics for anomaly detection with a focus on security analytics, network alerts, temporal characterization, time series analysis, and anomaly detection. Pierazzi et al. (2016) referred to simulation modeling and analysis, time series modeling, exploratory data analysis, and descriptive modeling as part of their work. To achieve their purpose, Pierazzi et al. (2016) proposed a framework for the automatic investigation of security alerts to determine whether and which anomaly detection algorithms can be applied effectively. As a result, Pierazzi et al.’s (2016) proposed framework is able to extract relevant descriptive statistics that allow analysts to understand the effectiveness of popular anomaly detection approaches with respect to different alerts groups. Pierazzi et al. (2016) found that although alerts exhibit different behaviors and statistics, the framework can evaluate the “effectiveness of the most popular anomaly detection approaches even in dynamic contexts influenced by many endogenous and exogenous factors that determine high variations in the number and nature of security alerts” (Pierazzi et al., 2016, p. 46). Pierazzi et al. (2016) proposed additional research focused on the integration of studies on cross-correlation and causality of alerts series, as well as solutions for online tuning and estimation of parameters specific to anomaly and system state change detection. Such framework tactics are vital to improving detection capabilities, particularly where anomaly detection via machine learning models can improve security event detection and alert quality.

A support vector machine (SVM) intrusion detection model, based on compressed sampling, was proposed by Chen et al. (2016) with attention to intrusion detection, efficiency, data sampling, and compression via compressed sensing theory. Chen et al. (2016) intended to reduce data dimensions in the data sampling stage and directly obtain the feature information of network data to improve the efficiency of security event detection. Chen et al.’s (2016) intrusion detection model, relative to direct use classifiers for learning, and detection of training and testing sets, had no significant change of security event detection rates and false positive rates. Chen et al. (2016) did however find that training and detection time was greatly reduced, which is key to improved detection in network data flows. Chen et al. (2016) concluded that this method can realize detection of network anomaly behavior quickly without reducing the classification accuracy. Chen et al. (2016) asserted that many network datasets need rapid and real-time detection, and that intrusion detection based on compressed sensing provides a real-time network security protection mechanism. Compression and sampling, and successful analysis of such data sets, are foundational to improved efficiency and the ability to actively process large data volumes for security analyst action.

Veeramachaneni et al. (2016) trained a big data machine to defend with a focus on problems specific to big data, behavioral analytics, and supervised learning. Veeramachaneni et al. (2016) intended to combine analyst intuition with state-of-the-art machine learning to build an end-to-end active learning system using a “big data behavioral analytics platform, an ensemble of outlier detection methods, a mechanism to obtain feedback from security analysts, and a supervised learning module” (Veeramachaneni et al., 2016, p. 1). As a result, Veeramachaneni et al. (2016) found that detection rates improved by an average of 3.41 times, false positives were reduced fivefold, and that the system was capable of learning to defend against unseen attacks. Veeramachaneni et al. (2016) concluded that the “system learns to defend against unseen attacks: as time progresses and feedback is collected, the detection rates increase” (p. 12). The findings imply that the Veeramachaneni et al. (2016) system achieved a detection rate of 86.8% even at an extremely low daily investigative budget of 200 events, which is a tenfold improvement over the unsupervised outlier detection approach rate of 7.9%. Veeramachaneni et al.’s (2016) research offers direct and immediate implications for this research in that it directly addresses and validates the use of data science/machine learning, coupled with visualization of resulting data outcomes, to aid the analyst in improved identification and response, thus increasing efficacy and satisfaction.

Huang (2015) utilized scoring algorithms to calculate security analysts’ performance scores according to weighted functions combing all performance metrics, as pertinent to situational awareness (SA) for security analysts in an information fusion environment. Huang (2015) set out to teach and measure individual and team situation awareness in a cyber-defense context and incorporate various technologies to enhance cyber analysts’ learning process. Huang (2015) found that the system’s ability to take high false positive data as input, and filter out only true positives, makes many sources of raw data, previously unusable, become reliable detectors of Advanced Persistent Threats (APTs). Huang (2015) asserted that this change allows a corresponding decrease in false negatives, meaning that more APTs will be detected. Huang (2015) utilized simulation in a cloud-based testbed and demonstrated that the designed system could effectively detect important types of APT behaviors including slow port scan attacks, slow brute force attacks, and data exfiltration attacks. Huang (2015) also established significantly reduced false positives as reported to the security analysts. Huang (2015) set out to build a real-time network status visualization GUI for the APT detection system. The visualization engine periodically checks the database for the outputs/updates of the corresponding analytic services, and then displays/updates the events on the topology and swim-lane views, allowing analysts to gain additional contextual knowledge (Huang, 2015).

The applicability of these methods remains paramount. Specialist security analysis, particularly the proactive hunt for adversaries and adversarial behavior, requires a combination of practical knowledge in programming, visualization, and data science (Slatman, 2016). As an example, Pierazzi et al.’s (2016) framework investigates temporal trends and patterns in security alerts, and automatically extracts relevant descriptive statistics used to understand the applicability of popular anomaly detection methods rooted in data science. Pierazzi et al. (2016) achieved this by separating security alerts into different groups and analyzing their data distribution and temporal dependence at different times. The descriptive statistical output is useful to acquire information pertinent to the security status of observed systems (Pierazzi et al., 2016). Vinayakumar et al. (2019) utilized deep neural network (DNN), a deep learning model, to develop their flexible and effective intrusion detection system (IDS) to detect and classify unforeseen and unpredictable cyberattacks. Vinayakumar et al. (2019) do so by combining network intrusion detection (NIDS) and host intrusion detection (HIDS) collaboratively with deep neural network (DNN) models to detect cyberattacks proactively to determine whether network traffic behavior is either normal or abnormal by classifying attack patterns into corresponding attack categories. The resulting model output is then subject to further analysis and visualization.

### Security Data Visualization

Visualization as a tool and mechanism for enhanced security data analysis serves many purposes. Security alerts can be classified and displayed in a visualization manner where persistent activity that raises alerts of the same type over a long period can be statistically filtered to reduce false alarm rates and support further visualization (Chivers et al., 2010). Amar et al., (2005) indicated that a data-centric view of information visualization relies on user skill to generate insight. Effective representation is a prerequisite to useful visualization. A focus on closely mapping visualization systems to user analytic goals increases the value and utility of information visualization (Amar et al., 2005). Information visualization benefits from understanding tasks that users undertake during actual analytic activity, where this understanding aids in creating visualizations that amplify users’ analytic abilities (Amar et al., 2005). Learning in a domain such as security data analysis benefits from the use of properly structured visualizations where interesting relationships between domain parameters can be discovered, a central tenet of security data analysis (Amar et al., 2005). These relationships are often denoted visually as graphs or networks. Security data visualization gives a meaningful overview of graph structure to highlight central objects, show similar objects, and reveal outliers (Vehlow et al., 2017).

Unique aspects of cyber-attacks, including information overload and rapid pace, fatigue, time, pressure, and anxiety challenge the security analyst. Security data visualization can be used to address related human cognitive limitations (Rajivan, 2011). Giacobe (2013) discussed the effectiveness of visual analytics and data fusion techniques on situation awareness in cyber-security, and focused on visual analytics, data fusion, and cybersecurity. Giacobe (2013) found that participants using the visual analytics (VA) interface performed better than those on the text-oriented interface, where the visual analytic interface yielded a performance that was quicker and more accurate that the text interface (Giacobe, 2013). Giacobe (2013) discovered that the benefits of visual analytics far outweighed text only. He did so utilizing a measurement technique to evaluate the impact of new visualizationsand fusion tools on security analyst situational awareness (SA). Giacobe (2013) also assessed the design, conduct and analysis of human-in-the-loop experiments to measure situational awareness using multiple situational awareness assessment instruments. Giacobe’s (2013) findings have implications for situational awareness in the cyber-security domain and for situational awareness and decision making in all domains using visual analytics as opposed to text-only analytics. This is paramount research for this study, as Giacobe (2013) has explored related themes specific to the value of visualization.

Giacobe (2013) also established that a picture is worth a thousand alerts, with a continued emphasis on visual analytics, data fusion, and cybersecurity. Giacobe (2013) asserted that to better understand a phenomenon such as situational awareness (SA), a multi-method approach of cognitive assessment may be useful, especially where tasks are complex and virtual such as those in cyber-security. Giacobe’s (2013) purpose for this research was to measure situation awareness using multiple techniques in a complex cyber-security task. To do so, Giacobe (2013) utilized a 2x2 between-subjects experimental design, constructed using two interfaces (text and visual analytic) and two groups of participants (novice and experienced). As a result of this experiment, where Giacobe (2013) increased the simulated task fidelity to measure situation awareness using different methods with participants of differing experience levels, he again determined that the visual analytic interface yielded quicker and more accurate performance than the text interface. Giacobe’s (2013) research implies that a layperson’s consideration for good situation awareness is a combination of phenomena, including performance, specific situation knowledge, low perceived workload, and a perception of effectiveness (Giacobe, 2013).

Visualization techniques have been developed for the exploration of dynamic networks for patterns and potential abnormal behaviors (Bender-de-Moll & McFarland, 2006). Two typical visualization methods for dynamic network analysis include small multiples and animation (van den Elzen et al., 2014; van den Elzen et al., 2016). If visualization is the cornerstone for situation awareness in a security analysis context, security analysts must be able to interact with the data and information on strategic, operational, and tactical levels (Slatman, 2016). Visualization techniques are useful for dynamic network analytic processes where the ability to comprehensively display spatiotemporal network information, control visual complexity, and improve computational scalability all benefit the security analyst (Liao et al., 2018).

Marty (2008) stated that "a picture is worth a thousand log records” (p.2). While Marty's (2008) book is dated in terms of current technology and capability, including data science and ML, the spirit and tenor of its message remains consistent. A visual representation of log records communicates the content of the log and allows viewers to process information in a fraction of time compared to reading the original log (Marty, 2008). Security visualization can conceptually be reduced to the process of generating a picture based on log records, and how log records are mapped to visual representations (Marty, 2008). Marty (2008) further contends that visual representations of data enable communication of large amounts of data and information to viewers rather than in text where it is more difficult to immediately comprehend. In general, it is harder for humans to process text, while images can be processed extremely well (Marty, 2008). Marty (2008) asserted that visualization can encode a wealth of information and communicate much larger amounts of data to viewers than non-visual methods. Visualizations can utilize shape, color, size, relative positioning, and other parameters to encode information, increasing bandwidth between the information and viewer (Marty, 2008). Visualization benefits are many, including the improved ability to answer questions, posing new questions, supporting decisions, communicating information, and increasing efficiency (Marty, 2008). More specifically, security visualization is driven by the increasing amount of data collected in computing environments, the security analyst's requirement for event and log analysis, regulatory compliance, and cybercrime prevention (Marty, 2008). Marty (2008) described a security visualization dichotomy driven by two issues: many security visualization tools are written by security people with little experience in visualization theory and human-computer interaction, while the rest are written by visualization people with little computer security and IT experience.

Jacobs and Rudis (2014) took their readers on a journey into the world of security data science with a focus on Python and R as foundational data analysis tools while introducing the design and creation of modern static and interactive visualizations. Jacobs and Rudis (2014) asserted that the ability to perform good data analysis and produce informative visualizations is a means to an end. The analysis of data and creation of visualizations are useful to gain new perspectives, to find relationships not yet discovered, or to uncover new information in the name of learning (Jacobs & Rudis, 2014). The realm of cybersecurity is vast, with many components, and massive complexity; analyst intuition must be augmented and supported with the science of data analysis (Jacobs & Rudis, 2014). Jacobs and Rudis (2014) focused on using real data, with more attention on process and less on results. The use cases are exemplary and teach new ways to look at and learn from data (Jacobs & Rudis, 2014).

In exploring why visualization of security data is important, Jacobs and Rudis (2014) asserted that the most efficient path to human understanding is through the visual sense. For humans to learn about systems and understand how they function (or do not function), they must exploit visual modalities to achieve their goal (Jacobs & Rudis, 2014). This goal is often to communicate points of interest effectively and efficiently from data where data visualization offers many advantages as a communication tool compared to other methods (Jacobs & Rudis, 2014). Data visualization offers these advantages:

* **Data visualizations communicate complexity quickly.** with visualization, millions of data points can be communicated in seconds while minimizing the loss of detail and resolution.
* **Data visualizations enable recognition of latent patterns.** patterns not apparent via statistical methods or data scanning may surface through visualization, and when presented visually, data patterns in a single variable, or relationships across many variables, may be illuminated.
* **Data visualizations enable quality control on the data.** mistakes and errors in data collection or preparation are revealed through visualization where data visualizations can serve to quality check work.
* **Data visualizations can serve as a muse.** laying out data visually can provide new perspective and facilitates the thinking and discovery processes (Jacobs & Rudis, 2014).

Jacobs and Rudis (2014) proposed that a performing data-driven security program within any organization should ensure that the program asks questions that have objective answers, finds, and collects relevant data, learns through iteration, and finds statistics. The preferred outcome is that the resulting data visualization efforts help consumers understand and learn from the data through effective communication delivered with effective visual creations (Jacobs & Rudis, 2014).

# Synthesis of the Research Findings

Recurring themes surface in exploration of the literature pertinent to the business problem at hand: organizations risk data breach, loss of valuable human resources, reputation, and revenue due to excessive security alert volume and a lack of fidelity in security event data. To illuminate the lifecycle of the information security analyst charged with analyzing and responding to security events and alerts, the literature review considers the related challenges these analysts and their organizations face via exploration of numerous perspectives.

Table 1. Security Analyst Workload Detail

|  |  |
| --- | --- |
| Security Information and Event Management (SIEM) | Platforms and practice specific to security event and alert processes. |
| Security Analyst Challenges:  Scale, Volume, Noise, and Workload | SIEMS can generate challenges for analysts, represented in issues of data scale, alert volumes, signal to noise ratio, and excess workloads. |
| Security Data Analysis and Processing | The methods by which these issues and challenges are addressed. |
| Machine Learning (ML) and Data Science (DS) | Security data analysis and processing methods can be greatly accentuated and optimized with the use of machine learning and data science. |
| Security Data Visualization | Output of these accentuated and optimized methods (models) can be represented via the practice of security data visualization. |

Table 2. Security Analyst Workload Cycle

The efficacy and expertise of the security analyst depends on their ability to tune their cognitive skills, including sustained attention and information synthesis, to security data, events, and alerts, in contextualized ways to access and deploy detection and response. Therefore, the security analyst must be armed with the ability to do so with speed and at scale (Ben-Asher & Gonzalez, 2015). As noted in the literature review, there are theories, methods, and practices that address the challenges and issue confronting the security analyst, yet attention to usability and ease of use must be applied. While not directly relevant to this study, a brief discussion is warranted specific to decision field theory as an applicable factor in the workload that the contemporary security analyst faces. As such, decision field theory can aid in the synthesis of the workload cycle of the security analyst and highlights the importance of viable solutions with output that compliments ease of use and intention to use. Decision field theory encompasses part of the psychological experience for the security analyst in that it is a fusion of approach-avoidance theories of motivation and information-processing theories of choice response time (Busemeyer & Townsend, 1993). Security analysts are faced with difficult and burdensome decisions in the face of large security data volumes, incomplete data, false positive alerts, and varied levels of severity. Approach-avoidance theories of motivation come to bear due to the coping methods security analysts must apply when confronted with overwhelming workload or data input. How then does the security analyst cope? When assessing the impact of information technology threat avoidance factors on avoidance behavior of users, Rho and Yu (2011) defined coping as cognitive and behavioral effort to manage internal/external requirements evaluated as something that occupies or overwhelms the resources of an individual. This is certainly a state the security analyst as described in the literature review can find themselves in on a regular basis, resulting in a scenario where the analyst might become less effective or focused. Approach-avoidance motivation includes considerations of severity. Carpenter et al. (2019) hypothesized that when individuals believe they are susceptible to a given technology threat, they will become more concerned about the severity of the consequences that result from that threat. This could again lead to a distracted or unfocused security analyst becoming worried about the consequence of failing to escalate a security alert properly or missing a true positive, and therefore becoming even less effective. Conversely, when that same security analyst knows that they are *not* susceptible to a technology threat, their awareness, concern, or fear of that threat, and the severity of its consequences, decreases. Thus, means and methods to reduce this sense of susceptibility are critical to increased security analyst success. This study will seek to validate the usability of, and intent to use of, refined output from advanced methods to improve the security analyst’s performance.

Information-processing theories specific to choice response time are applicable as part of decision field theory. Again, when considering the plethora security analyst workloads, each including issues of scale, volume, noise, security data analysis and processing, choice response time is inherent throughout. Where choice response time is the aggregate of decision making and reaction time, tactics are needed to provide decision support. Giacobe (2013) focused on this topic in his study of situation awareness for security analysts. The challenge Giacobe (2013) sought to address was security analysts’ perception of their mental model as appropriate for their task at hand. Further, Giacobe (2013) researched if the decisions being made were appropriate specific to response time, decision making and reaction time. Giacobe (2013) determined that security data fusion methods combined security data output to support improved inference and decisions made by security analysts. These techniques can work regardless of uncertainty and can help synthesize high security data volumes into meaningful information to help security analysts make decisions (Giacobe, 2013). Fundamentally, again, the question at hand is this. Is there a difference between the adoption of visual alert output (VAO) and text alert output (TAO) as predicted by the Technology Acceptance Model (TAM)?

# Critique of Previous Research Methods

The literature reviewed is consistent in its academic rigor and theoretical support, but recurring limitations surface. While academic research is critical to the betterment of the field, challenges arise in the actual applicability of recommendations and results in a working production environment or commercial venue. Another general criticism of the related literature is specific to the different sub-domains of cyber-security, where specific measurement techniques and instruments may not be globally applicable to all sub domains (Giacobe, 2013). Ideally, the methodology of developing specific instruments should be generalized to different sub-domains (Giacobe, 2013).

Anthropological studies focused on security analyst burnout due to excess workload and security data volumes could only validate the proposed model for a solution at one research site, specifically, a security operations center (SOC; Sundaramurthy et al., 2015). Additionally, under similar circumstances, Sundaramurthy et al. (2015) could not declare their model as exhaustive due to a limited 6-month research period that could not factor for new events that might occur after the conclusion of fieldwork. Where research may lead to effective and efficient design of information security systems to better support the work of the security analyst, it must factor for the capabilities and limitations of the human decision maker who uses them (usability; Kaufman, Perlman, & Speciner, 2002). An understanding of how past experiences and knowledge influence decision making in a highly dynamic realm such as security data analysis is vital; not all research accounts for it (Ben-Asher & Gonzalez, 2015).

Research limitations specific to available security datasets are a reality. Many datasets representing current network attacks are kept private due to privacy and operational security concerns (Vinayakumar et al., 2019). On the other hand, publicly available datasets are almost always anonymized and thus lack validity as real-world network traffic profiles (Vinayakumar et al., 2019).

Certain well researched solutions intended to ease the security analyst workload can compound the issues. As part of security analysis, where data provenance, the reconstruction of the chain of events leading to an alert, is often considered valuable, it is both labor intensive and creates additional dependencies, leading to a higher potential for false positive alerts (Hassan et al., 2019). To achieve improved security analysis to predict stages of a multi-state cyber-attack, Yang et al. (2009) designed an alert fusion engine comprised of an Information Fusion Engine for Real-time Decision-making (INFERD) and Threat Assessment for Network Data and Information (TANDI). INFERD performs intrusion detection system (IDS) alert correlation while TANDI determines threats to the network based on the aggregation of output from INFERD and network configurations (Lyons, 2014). Yet, in his research, Giacobe (2013) found that while INFERD and TANDI showed promising results, the threat assessment and prediction algorithms are based on a priori information and thus are susceptible to the knowledge limitations of the security analysts providing the information. This is in keeping with Ben-Asher and Gonzalez (2015) who concluded that expertise and practical knowledge play an important role in security alert triage analysis.

Though various machine learning based solutions are found in the literature, the applicability to commercial systems is at an early stage (Vinayakumar et al., 2019). Additionally, certain machine learning techniques suffer from high-false positive detection due to limitations in differentiating attack behavior and evolving normal behavior (Mishra et al., 2019). In summary, it is imperative to understand what machine learning models are doing, evaluate the models, reduce the scope of models, and understand what the limitations of machine learning models are when deploying machine learning in production information security settings (Sommer & Paxson, 2010). The use of algorithms in the cyber security domain is less about invention and more about selection (Giacobe, 2013). Thus, Giacobe (2013) asserted that, while high data volumes specific to cybersecurity pose unique problems and limitations, the selection of the appropriate algorithm for use in security data processing can be enhanced when fused with successes from other cybersecurity domains.

Even in the realm of security data visualizations there are shortcomings. Aspirational techniques such as mVis (multivariate Visualiser) include interactive visual interfaces for data exploration providing meaningful selection and labelling of records based on insight determined by the user (Chegini et al., 2019). Yet, Chegini et al. (2019) stated that mVis has visualization and algorithmic limitations where the visual scalability of the label alphabet (number of partitions) is limited to twelve distinct colors, no more colors can be comfortably distinguished. Additionally, the scatterplot matrix (SPLOM) and parallel coordinates views are limited by the amount of available screen space (Chegini et al., 2019). Where Liao et al. (2018) referred to visualization techniques developed for the exploration of dynamic networks for patterns and potential abnormal behaviors (Bender-de-Moll & McFarland, 2006), they asserted that it is still difficult to track changes over time due to limitations of human cognitive ability. Other three-dimensional (3D) dynamic network visualizations that superimpose and/or layer suffer from issues with visual clutter, readability, and preservation of a mental map (Liao et al., 2018). Again, usability and ease of use surface as requirements for such security data outputs.

# Summary

The challenges security analysts must overcome to respond appropriately to threats and attacks are many. The uncertainty of attacker behaviors, the complexity of interconnections between resources in modern distributed systems, huge security datasets, time limitations, and balancing between countermeasure costs and attack losses are all factors (Kotenko & Doynikova, 2017). Organizations can expect systemic challenges as well, specific to their information security, each of which can be addressed with data science, machine learning, and security data visualization (Veeramachaneni et al., 2016). A lack of labeled data, constantly evolving attacks, and limited investigative time and budget all demand a different approach for organizations to address the business problem (Veeramachaneni et al., 2016). To properly address these challenges, security analysts’ time must be used effectively, with mechanisms to detect new and evolving attacks in early stages, reduce response time between security event detection and attack prevention, and ensure an acceptably low false positive rate (Veeramachaneni et al., 2016). Of critical importance to individual security analysts, and their respective organizations, is the perceived usefulness and ease of use of the output generated by the aforementioned mechanisms to address the challenges explored in this literature review. The resulting output from data science and machine learning methods, be it text based, or security data visualized, is the consequence of numerous factors (Hassenzahl & Tractinsky, 2006). The security analyst’s internal state (predispositions, expectations, needs, motivation, mood), the characteristics of the resulting output (complexity, purpose, usability, functionality), and the context within which the interaction occurs (organizational/social setting, meaningfulness of the activity, voluntariness of use) all come to bear (Hassenzahl & Tractinsky, 2006). Thus, user acceptance of the DS and ML output is an imperative, regardless of text or visual orientation. Perceived usefulness, perceived ease of use, and intention to use must also be met with improvements for the security analyst’s experience, including speedier task completion, improved job performance, increased productivity, enhanced effectiveness, and general job improvement for the security analyst (Lewis, 2019). Should the output from a DS and ML solution be easy to use in a clear and understandable manner, it is more likely to be used regularly to greater satisfaction and success (Lewis, 2019). This study will endeavor to assess the perceived usefulness and perceived ease of use of the resulting security alert output of DS and ML models that are intended to solve security analysis challenges. The research will determine if there is a statistically measurable difference in the perceived usability of visualized alert output (VAO) and text-based alert output (TAO).

References

Amar, R., Eagan, J., & Stasko, J. (2005). Low-Level Components of Analytic Activity in Information Visualization. Paper presented at the IEEE Symposium on Information Visualization, Minneapolis, MN USA. https://doi.ieeecomputersociety.org/10.1109/INFOVIS.2005.24

Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. *Computers in Human Behavior, 48*, 51-61. doi:10.1016/j.chb.2015.01.039

Bender-deMoll, S., & McFarland, D. A. (2006). The Art and Science of Dynamic Network Visualization. *J. Soc. Struct., 7*.

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. doi:10.1080/10447318.2015.1064648

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. Psychological Review, 100(3), 432-459. doi:10.1037/0033-295x.100.3.432

Carpenter, D., Young, D. K., Barrett, P., & McLeod, A. J. (2019). Refining Technology Threat Avoidance Theory. *Communications of the Association for Information Systems*, 380-407. doi:10.17705/1cais.04422

Chegini, M., Bernard, J., Berger, P., Sourin, A., Andrews, K., & Schreck, T. (2019). Interactive labelling of a multivariate dataset for supervised machine learning using linked visualisations, clustering, and active learning. *Visual Informatics, 3*(1), 9-17. doi:10.1016/j.visinf.2019.03.002

Chivers, H., et al. (2010). Knowing who to watch: Identifying attackers whose actions are hidden within false alarms and background noise. *Information Systems Frontiers, 15*(1): 17-34.

Cisco (2017). Cisco 2017 Annual Cybersecurity Report. San Jose, CA, Cisco.

CriticalStart. (2019). *The Impact Of Security Alert Overload*. Retrieved from Plano, TX: https://www.criticalstart.com/

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly, 13*, 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace. Journal of Applied Social Psychology, 22(14), 1111–1132. [https://doi.org/10.1111/j.1559-1816.1992.tb00945.x](https://psycnet.apa.org/doi/10.1111/j.1559-1816.1992.tb00945.x)

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies, 45*(1), 19-45. doi:https://doi.org/10.1006/ijhc.1996.0040

Dimensional Research. (2020). *2020 State of SecOps and Automation*. Retrieved from https://www.sumologic.com/brief/state-of-secops

Dutt, V., Ahn, Y., & Gonzalez, C. (2013). Cyber situation awareness. *Human Factors: The Journal of Human Factors and Ergonomics Society, 55*(3), 605-618. doi:10.1177/0018720812464045

FireEye. (2015). *The numbers game: How many alerts is too many to handle?* Retrieved from https://www.fireeye.com/offers/rpt-idc-numbers-game-special-report.html

Giacobe, N. A. (2013). A picture is worth a thousand alerts. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 172-176. doi:10.1177/1541931213571039

Giacobe, N. A., McNeese, M. D., Mancuso, V. F., & Minotra, D. (2013). Capturing human cognition in cyber-security simulations with NETS. In IEEE ISI 2013 - 2013 IEEE International Conference on Intelligence and Security Informatics: Big Data, Emergent Threats, and Decision-Making in Security Informatics (pp. 284-288). [6578844] (IEEE ISI 2013 - 2013 IEEE International Conference on Intelligence and Security Informatics: Big Data, Emergent Threats, and Decision-Making in Security Informatics). <https://doi.org/10.1109/ISI.2013.6578844>

Giacobe, N. A. (2013). *Measuring the effectiveness of visual analytics and data fusion techniques on situation awareness in cyber-security* (Doctoral dissertation). Retrieved from Penn State Electronic Theses and Dissertations. 17537

Greenlee, E. T., et al. (2016). Stress and Workload Profiles of Network Analysis: Not All Tasks Are Created Equal. *Advances in Human Factors in Cybersecurity***,** 153-166.

Grier, RA 2015, ‘How high is high? A meta-analysis of NASA-TLX global workload scores’,

*Human Factors and Ergonomics Society Annual Meeting: Proceedings*, *59*(1), pp. 1727-

31.

Hassan, W. U., Guoz, S., Li, D., Chen, Z., Jee, K., Li, Z., & Bates, A. (2019). NODOZE: Combatting threat alert fatigue with automated provenance triage. Paper presented at the 2019 Network and Distributed System Security Symposium. doi:10.14722/ndss.2019.23349

Hassenzahl, M. (2005). The thing and I: Understanding the relationship between user and product. In M. Blythe, K. Overbeeke, A. Monk, & P. Wright (Eds.), Funology: From usability to enjoyment (Vol. 3, pp. 31–42). Berlin, Germany: Springer. doi:10.1007/1-4020-2967-5\_4

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology, 25*(2), 91-97. doi:10.1080/01449290500330331

Henderson, R., & Divett, M. J. (2003). Perceived usefulness, ease of use and electronic supermarket use. *International Journal of Human-Computer Studies, 59*(3), 383-395. doi:10.1016/s1071-5819(03)00079-x

Huang, Z. (2015). *Human-centric training and assessment for cyber situation awareness* Available from Dissertations & Theses Europe Full Text: Science & Technology. Retrieved from <https://search.proquest.com/docview/1767223484>

Igbaria, M., Parasuraman, S., & Baroudi, J. J. (1996). A motivational model of microcomputer usage: JMIS JMIS. *Journal of Management Information Systems, 13*(1), 127.

Jacobs, J., & Rudis, B. (2014). Data-Driven Security: Analysis, Visualization and Dashboards (C. Long Ed.). Indianapolis, IN: John Wiley & Sons, Inc.

Kaufman, C., Perlman, R., & Speciner, M. (2002). *Network security: Private communication in a public world*. Englewood Cliffs: Prentice Hall.

Kavanagh, K., Bussa, T., & Sadowski, G. (2020). Magic Quadrant for Security Information and Event Management (G00381093). Retrieved from Gartner website: https://www.gartner.com/doc/reprints?id=1-1YE69EYM&ct=200218&st=sb

Khalili, M., Zhang, M., Borbor, D., Wang, L., Scarabeo, N., & Zamor, M.-A. (2019). Monitoring and Improving Managed Security Services inside a Security Operation Center. *ICST Transactions on Security and Safety, 5*(18). doi:10.4108/eai.8-4-2019.157413

Kotenko, I. and E. Doynikova (2017). "Selection of countermeasures against network attacks based on dynamical calculation of security metrics." *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology 15*(2): 181-204.

Lane, D. M. (2021). *Online Statistics Education: A Multimedia Course of Study*. Retrieved from https://onlinestatbook.com/

Legris, P., Ingham, J., & Collerette, P. (2001). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management, 40*(2003), 191-204.

Lewis, J. R. (2019). "Comparison of Four TAM Item Formats: Effect of Response Option Labels and Order." *Journal of Usability Studies 14*(4): 224-236.

Lewis, J. R., Utesch, B. S., & Maher, D. E. (2015). Measuring Perceived Usability: The SUS, UMUX-LITE, and AltUsability. *International journal of human-computer interaction, 31*(8), 496-505. doi:10.1080/10447318.2015.1064654

Liao, Q., Li, T., & Blakely, B. A. (2018). Anomaly analysis and visualization for dynamic networks through spatiotemporal graph segmentations. *Journal of Network and Computer Applications, 124*, 63-79. doi:10.1016/j.jnca.2018.09.016

Lynn, S. K. and L. F. Barrett (2014). ""Utilizing" signal detection theory." *Psychological Science 25*(9), 1663-1673.

Lyons, K. (2014). *A Recommender System In The Cyber Defense Domain.* (Master of Science). Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio.

Marty, R. (2008). *Applied Security Visualization* (K. Gettman Ed.). Boston, MA: Addison-Wesley.

Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2019). A Detailed Investigation and Analysis of Using Machine Learning Techniques for Intrusion Detection. *IEEE Communications Surveys & Tutorials, 21*(1), 686-728. doi:10.1109/comst.2018.2847722

Moustafa, N., Hu, J., & Slay, J. (2019). A holistic review of network anomaly detection systems: A comprehensive survey. *Journal of Network and Computer Applications, 128*, 33-55. doi:10.1016/j.jnca.2018.12.006

Nayak, M. S., & Narayan, K. A. (2019). Strengths and Weakness of Online Surveys. *IOSR Journal Of Humanities And Social Science, 24*(5), 31-38. doi:10.9790/0837-2405053138

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Journal of Educational Technology & Society, 12*(3), 150-162.

Paul, C. L., & Dykstra, J. (2017). Understanding operator fatigue, frustration, and cognitive workload in tactical cybersecurity operations. *Journal of Information Warfare, 16*(2), 1-11. Retrieved from www.jinfowar.com

Pierazzi, F., et al. (2016). "Exploratory security analytics for anomaly detection." *Computers & Security, 56*: 28-49.

Ponemon. (2017). *The Cost of Insecure Endpoints*. Retrieved from https://www.absolute.com/resources/analyst-reports/ponemon-the-cost-of-insecure-endpoints/

Ponemon. (2020). *2020 Devo SOC Performance Report: A Tale of Two SOCs*. Retrieved from Cambridge, MA: https://www.devo.com/

Rajivan, P. (2011). CyberCog, Arizona State University. Master of Computing Studies: 74.

Rho, H., & Yu, I. (2011). The impact of information technology threat avoidance factors on avoidance behavior of user. Retrieved from https://pdfs.semanticscholar.org/a9e9/31b792c403f49a91a1dbdfdfbc34b92f1d8c.pdf

Sawyer, B. D., et al. (2014). "Cyber Vigilance: Effects of Signal Probability and Event Rate." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*(1): 1771-1775.

Slatman, H. (2016). Unboxing Security Analytics: Towards Effective Data Driven Security Operations. Computer Science, University of Twente.

Sommer, R., & Paxson, V. (2010, 16-19 May 2010). *Outside the Closed World: On Using Machine Learning for Network Intrusion Detection.* Paper presented at the 2010 IEEE Symposium on Security and Privacy.

Sundaramurthy, S. C., Bardas, A. G., Case, J., Ou, X., Wesch, M., McHugh, J., & Rajagopalan, S. R. (2015). *A Human Capital Model for Mitigating Security Analyst Burnout.* Paper presented at the SOUPS.

van der Heijden, H. (2004). User Acceptance Of Hedonic Information Systems . *MIS Quarterly, 28*(4), 695-703.

Veeramachaneni, K., Arnaldo, I., Korrapati, V., Bassias, C., & Li, K. (2016). AI^2: Training a Big Data Machine to Defend. 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS). doi:10.1109/bigdatasecurity-hpsc-ids.2016.79

Vehlow, C., Beck, F., & Weiskopf, D. (2017). Visualizing Group Structures in Graphs: A Survey. *Computer Graphics Forum, 36*(6), 201-225. doi:10.1111/cgf.12872

Venkatesh, V., & Davis, D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Information Systems Research, 46*(2), 186–204.

Vieane, A., Funke, G., Gutzwiller, R., Mancuso, V., Sawyer, B., & Wickens, C. (2016). Addressing human factors gaps in cyber defense. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 770-773. doi:10.1177/1541931213601176

Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep Learning Approach for Intelligent Intrusion Detection System. *IEEE Access, 7*, 41525-41550. doi:10.1109/ACCESS.2019.2895334

Wu, H. and Z. Wang (2018). "Multi-source fusion-based security detection method for heterogeneous networks." *Computers & Security, 74*, 55-70.

Wu, J., Chen, Y., & Lin, L. (2007). Empirical evaluation of the revised end user computing acceptance model. Computers in Human Behavior, 23(1), 162-174. doi:10.1016/j.chb.2004.04.003

Wu, J.-H., & Wang, S.-C. (2005). What drives mobile commerce? *Information & Management, 42*(5), 719-729. doi:10.1016/j.im.2004.07.001

Yang, S., Stotz, A., Holsopple, J., Sudit, M., & Kuhl, M. (2009). High level information fusion for tracking and projection of multistage cyber attacks. *Information Fusion, 10(1)*, 107-121. doi:10.1016/j.inffus.2007.06.00