Analysis of the Design Space for Cybersecurity Visualizations in VizSec

Adrian Komadina * Željka Mihajlović † Stjepan Groš ‡

Faculty of Electrical Engineering and Computing, University of Zagreb
Department of Electronics, Microelectronics, Computer and Intelligent Systems
Laboratory for Information Security and Privacy
Zagreb, Croatia

ABSTRACT

In this paper, we present research on the analysis of the design space for cybersecurity visualizations in VizSec. At the beginning of this research, we analyzed 17 survey papers in the field of cybersecurity visualization. Based on the analysis of the focus areas in each of these survey papers, we identified five key components of visualization design, i.e. *Input Data*, *Security Tasks*, *Visual Encoding*, *Interactivity*, and *Evaluation*. To show how research papers align with these components, we analyzed 60 papers published at the *IEEE Symposium on Visualization for Cyber Security (VizSec)* between 2016 and 2021 in the context of the five identified components. As a result, each research paper was classified into several categories derived from the selected components of the visualization design. Our contributions are: (i) an analysis of the focus areas in survey papers on cybersecurity visualization and (ii) the classification of 60 research papers in the context of the selected components of the visualization design. Finally, we highlighted the main findings of the analysis and drew conclusions.

Index Terms: Visualization—Cybersecurity—Analysis—Survey VizSec

1 INTRODUCTION

Visualization not only provides a beautiful and attractive representation of a dataset, but also plays an increasingly important role in many sciences. Especially in the field of data science, visualization has become one of the most important tools in recent years. Visualization has many benefits, but the most important are answering questions about a dataset quickly, raising new questions, exploring a dataset and gaining new insights, making decisions faster and more accurately, increasing the efficiency of data analysis, and encouraging users to try new ways of visualization [52]. Moreover, visualization enables users to identify concepts and relationships in datasets by explicitly displaying properties and relationships that are inherent and implicit [72].

In cybersecurity, we often talk about information visualization. In this context, the information we need to represent is often encoded in textual form. When information is encoded in this way, it is difficult for the human brain to process it, especially when the amount of data is large. On the other hand, if such data is represented as an image, it can be easily processed. These representations help users quickly identify outliers, detect malicious activity, uncover misconfigurations and anomalies, and identify general trends and relationships between individual data points [52]. Recognizing patterns and anomalies provides the user with new knowledge and insights and encourages further investigation [72].

*e-mail: adrian.komadina@fer.hr
†e-mail: zeljka.mihajlovic@fer.hr
‡e-mail: stjepan.gros@fer.hr

Our analysis of survey papers we found showed that research papers in this area are scattered across many different conferences, but *The IEEE Symposium on Visualization for Cyber Security (VizSec)* [80] stands out as the most significant among them. VizSec Symposium addresses advances and new techniques in analytics and visualization to meet the needs and challenges of the cybersecurity community, with security and privacy as the main themes. Launched in early 2004 as *Workshop on Visualization and Data Mining for Computer Security* at *ACM CCS*, VizSec established itself as the only specialized symposium in the field of cybersecurity visualization. For this reason, we decided to focus on papers published at this symposium.

The main goal of this research was to perform an analysis of the design space in the created visualization systems from VizSec research papers. To achieve our goal, we managed to find 17 survey papers in the field of cybersecurity visualization and analyzed what they focus on. Based on this analysis, we identified five different components of visualization design for creating cybersecurity visualizations. These components consist of: the data used as input to the visualization, the security domain task to be solved by the visualization, the visual encoding chosen and the degree of interactivity used to accomplish the target task, and the evaluation of the visualization system created. In the context of these five selected components, we analyzed 60 VizSec research papers to provide an analysis of the design space in the visualization systems created by these papers. Each of these papers was categorized, with a new perspective on categorization within selected visualization design components.

The contributions of this work are:

- Analysis of 17 survey papers in visualization security domain and identification of the focuses of each of these surveys.
- Analysis and classification of 60 VizSec research papers from 2016 to 2021 based on the five selected components of visualization design.

This paper is organized as follows. Following this section, in Sect. 2 we described the methodology we used to find survey papers and discussed why we chose the VizSec Symposium as the primary source for the papers. In Sect. 3, we listed the surveys we found in the field and described the components of the visualization design they focus on. In Sect. 4, we analyzed and categorized 60 contributions to the VizSec Symposium based on the five selected components of visualization design. In Sect. 5 we presented the main results of the analysis and some interesting correlations that emerged from the relationships between the observed categories. Related work is presented in Sect. 6. Finally, in Sect. 7 we presented the conclusion of the research with possible directions of future work.

2 METHODOLOGY OF THE RESEARCH

The way we collected the papers referenced in this article is as follows. The Google Scholar search engine was used to find the survey papers in this field. The keywords *visualization in security*, *security visualization*, *network security visualization*, *visualization*

in network security, *cybersecurity visualization*, and *VizSec* were used. Articles found as references to previously found work were also included. Based on these search, 17 survey papers were found that dealt with security visualization in some way.

Based on survey papers found, we analyzed where they were published and where their references were published. In this analysis, we found that while there was no overlap in the conferences where the surveys were published, but the references they contained were mostly from the VizSec Symposium.

Many references also come from the *IEEE Conference on Visual Analytics Science and Technology (VAST)* and the *IEEE Symposium on Information Visualization (InfoVis)*, but these mostly focus on the integration of machine learning and visualization. In addition, in 2021 VAST, InfoVis and *IEEE Scientific Visualization (SciVis)* have been combined into one *IEEE VIS: Visualization & Visual Analytics*, but is not considered in this paper due to lack of security topics.

From this it is clear that VizSec is considered the most important and lively conference in this field. For this reason, we decided to base our research on the contributions to VizSec. We limited ourselves to research papers from 2016 to 2021, which ultimately resulted in 60 papers, as we felt this was a representative set for our research. Papers dating back 6 years were easy to find using the IEEE Xplore digital library.

3 Analysis of Existing Surveys

The first part of this research is based on survey papers that are closely related to the topic of security visualization. During the research we were able to find 17 survey papers. We decided to categorize these papers according to the areas of interest they focus on within an observed domain.

Everything starts with data, including our categories. Input data is not only a close link between security and visualization, but also the basis for the choice of visualization methods. By this we mean not only the type of data itself, but also its processing and source. After selecting the data to be displayed, we encounter various obstacles, problems and challenges that come in our way. When we talk about the field of visualization, the following criteria arise from its essence, namely the visual encoding and interactivity realized in such a system. In order to validate the output system, it is necessary to perform an evaluation of the system. Finally, in order to justify the implementation of the whole solution, i.e. to give it a meaning, it is necessary to contextualize the created system by the type of use and the possible applications.

Based on the focus areas of selected surveys, we have established six categories: *Input Data*, *Tasks and Challenges*, *Visual Encoding*, *Interactivity*, *Evaluation*, and *Use cases and Applications*. In terms of visualization design taxonomy, these identified categories are fairly general when it comes to creating visualization systems, and can be found in similar form in other work. For example, Munzner [56] presents a four-level model for visualization design and evaluation. The first two levels of this model characterize the problems and data of a particular domain and map them into abstract operations and data types. The third level describes visual encoding and interactions. Finally, the fourth level deals with the creation of algorithm to implement this design. Munzner also includes evaluation as an integral part of this research to verify the validity of each level of the model.

In the Table 1 all surveys found are listed. For each paper, a check mark indicates which of the six selected areas of visualization design each survey focuses on.

From the table created, it is immediately apparent that the category most frequently observed in the surveys is the visual encoding itself. The most common way to look at the encoding itself is to look at the techniques used and their percentage of use in the visualization systems created over the years. Zhang et al. [89] divided visual designs into five categories based on the form of the visual

Table 1: List of the 17 survey papers categorized by the area of interest they focus on in cybersecurity visualization

Survey paper	Visual Encoding	Input Data	Use Cases and Applications	Tasks and Challenges	Evaluation	Interactivity
Shiravi et al. [72]	✓	✓	✓	✓		
Kasemsri [43]	✓	✓		✓		
Staheli et al. [76]					✓	
Ferebee and Dasgupta [29]	✓	✓				
Zhang et al. [89]	✓					
Muchagata and Ferreira [54]	✓					✓
Zhang et al. [88]	✓	✓	✓	✓		
Adams and Snider [1]	✓			✓		
Ji et al. [40]	✓	✓				
Kartel et al. [42]	✓	✓		✓		✓
Haina et al. [38]	✓		✓			
Kolomeec et al. [45]	✓	✓				
Liu and He [50]	✓	✓	✓		✓	
Crouser et al. [24]	✓		✓			
Tamassia et al. [79]	✓		✓			
Wagner et al. [86]	✓	✓		✓		✓
Langton and Baker [47]					✓	

results, while Kolomeec et al. [45] divided visualization approaches into two broad groups: Geometrical Models and Graphical Models. Ji et al. [40] also performed a classification of visualizations, identifying and exploring four key approaches for developing an effective network visualization system. Crouser et al. [24] took an alternative approach to examine existing VizSec publications using machine learning algorithms to identify common thematic groupings. When only graph drawing methods are considered, Tamassia et al. [79] categorized them by different areas of security visualization.

In addition to the visual encoding, the input data was also frequently observed. This is because the strongest reference to the cybersecurity field can be seen in the input data. The works that deal with input data can be roughly divided into two categories. One category includes works that look at the nature of the data itself, its origin, and its properties, and the other where data manipulations such as processing, filtering, and transformation are observed. Of the papers listed, Ferebee and Dasgupta [29] focused most on the input data through the data manipulations described. In terms of input data properties, Zhang et al. [88] classified network anomaly data based on the data type and data properties and shows the percentage of different visualization forms for each data property.

The next two categories in the table show how to create a visualization, i.e., requirements, constraints, challenges, and problems, and what to do with the finished product, i.e., how and where to use it. These two categories come up relatively frequently as areas of interest, but are rarely the focus of a work; rather, they are usually treated as extensions of another area. Shiravi et al. [72] classified security visualizations into five categories based on how they are used. In addition to use cases, Haina et al. [38] presented six applications for network security visualization. Kasemsri [43] categorizes the tasks of security visualization techniques into one of the following categories: finding intrusions, finding false alarms, and training classifiers. It also considers design issues related to high-dimensional data and multiple levels of detail. Adams and Snider [1] identified numerous challenges to the successful creation and implementation of visual tools for cybersecurity. Moreover, these two areas of interest are often combined into one, as which was made by Shiravi et al. [72] and Zhang et al. [88].

If we talk about evaluation, it was observed only in three works. Liu and He [50] indicated, among the works in the field of visualization in security, whether a work contains an evaluation or not, which is not very informative. Also, Langton and Baker [47] summarized a number of evaluation methods from the field of information visualization and shows how they can be applied in the field of cybersecurity. Staheli et al. [76] based the overview of evaluation on the papers from the VizSec Symposium. The papers found were categorized by the dimensions and components that are evaluated, as well as by the type of user and the type of evaluation, which provides a much more informative overview.

Interactivity is now an integral part of visualization, especially in the field of visual analysis. In the surveys found, it was observed three times. Muchagata and Ferreira [54] mentioned several types

of interaction related to information security analysis, considering the characteristics of each method. Kartel et al. [42] and Wagner et al. [86] showed only the proportion of works that use some form of interaction in the area of malicious code analysis. Although these three works formally address the issue of interactivity, there is no real analysis of it, which is the motivation for exploring it in this work.

From the analysis of these survey papers, we concluded that little research has been done on the categorization of evaluation and interactivity, while most work focuses on visualization techniques and input data. Of the six categories identified, we selected four: *Input Data*, *Visual Encoding*, *Interactivity*, and *Evaluation*. In addition to these four categories, we included an additional category *Security Tasks*. With this category, we associated the contributions with the corresponding security task they addressed using the created visualization system. This additional category can be seen as a smaller part of the two categories already identified, *Use Cases and Applications* and *Tasks and Challenges*, focusing more on the security aspect.

In the remainder of the paper, we analyze and categorize the VizSec Symposium research papers from the past six years in relation to the five selected areas of interest in the field of visualization, taking a fresh perspective on categorization within each area.

4 Analysis of the Components of Visualization Design

In this section, 60 papers from the VizSec Symposium are analyzed in the context of the five components of visualization design that emerged from the surveys in Sect. 3. Each of the five components is described in a separate subsection, and the VizSec contributions are classified based on various aspects for each component.

4.1 Input Data

We have categorized the input data used in the VizSec papers to provide better insight into the most commonly used data types. Data transformation and manipulation is not considered as in some other surveys, but whether the data source used is real or not.

The fidelity attribute of the data source provides information about whether the input data used to create a visualization system comes from the real world or is artificially generated (simulated). In some cases this attribute is not specified [21], and in others no dataset is used. These papers mostly use some specifications [8,71] or research papers [85] or even some other papers in the form of a survey [17]. For these papers, the fidelity attribute were not observed.

The Table 2 lists the 60 papers published at the VizSec Symposium in chronological order. The second column of the table contains the type of input data and the third column contains the fidelity attribute of the data source used. We can see that most papers use some kind of logs as input to their visualization system. In most cases, these logs are network traffic, DNS records or some kind of events. In addition to logs, malware samples and files in the form of executables or source files are also commonly used.

As for the fidelity attribute, we concluded that the vast majority of papers use a real-world data sources, about 82% compared to 18% of simulated ones. From our analysis, only Anh Huynh et al. [9] and Cappers and van Wijk [20] used combined real and simulated data sources. Simulated data is sometimes a stand-in for sensitive data types that cannot be used for experiments [41]. Most of the time, these data is personal data [23, 87], so it is obvious to use a simulated data source given the privacy and sensitivity of these data.

4.2 Security Tasks

Having successfully identified data inputs that are closely related to the security component of the papers, the next step is to examine their security context in more detail. We have classified the VizSec papers into 14 security themes based on the cybersecurity task that the visualization is intended to help with.

Table 2: 60 VizSec papers classified by type of input data and fidelity attribute of data source

Paper	Input data type	Fidelity
Sopan and Berlin [74]	Machine learning models, artifacts (files, URLs, emails)	Real
Gove [34]	Incident reports, network logs	Simulated
Angelini et al. [5]	Business functions, devices, vulnerabilities	Real
John et al. [41]	Sensitive data	Simulated
Graham et al. [36]	User behaviour logs, policy alerts	Real
Nadeem et al. [57]	Intrusion alerts	Real
Schreiber et al. [69]	Git repositories	Real
Reynolds et al. [64]	Binary code, vulnerabilities	Real
Böhm et al. [18]	System activities, file versions, network activity	Real
Dennig et al. [27]	Source code	Real
Schufrin et al. [70]	Personal data	Real
Beran et al. [16]	File metadata	Real
Alperin et al. [3]	Vulnerabilities, exposures	Real
Becker et al. [15]	DNS records	Real
Chaffey and Sgandurra [21]	Malware samples	N/A
Peng et al. [61]	Network traffic and events	Real
Guerra et al. [37]	Network traffic	Real
Laughlin et al. [48]	Machine learning dataset	Real
Varga et al. [85]	N/A	N/A
Subramanian et al. [77]	User behaviour logs	Real
Dasgupta et al. [26]	Machine learning dataset	Real
O'Shaughnessy [60]	Malware samples	Real
Ulmer et al. [84]	Network traffic	Simulated
Fouss et al. [30]	DNS records, network traffic	Real
Lohfink et al. [51]	Network sensor readings	Real
Angelini et al. [7]	Binary code	Real
Ošlejšek et al. [59]	Game events	Simulated
Chou et al. [23]	Personal data	Simulated
Sopan et al. [75]	Alerts	Real
Arendt et al. [11]	User activity events	Simulated
Cappers et al. [19]	Network traffic, malware samples	Real
Bakirtzis et al. [14]	Attack patterns, weaknesses, vulnerabilities	Real
Angelini et al. [6]	Program	Real
Yang et al. [87]	Personal data	Simulated
Chen et al. [22]	User behaviour logs	Real
Krokos et al. [46]	Network traffic	Real
Ulmer et al. [83]	Geo-IP data	Real
Gove and Deason [35]	DNS records	Real
Norton and Qi [58]	Machine learning dataset	Real
Angelini et al. [8]	N/A	N/A
Sethi and Wills [71]	N/A	N/A
Kim et al. [44]	Firewall rules	Real
Santhanam et al. [68]	Applications	Real
Theron et al. [81]	Network traffic	Real
Leichtnam et al. [49]	Network traffic	Real
Angelini et al. [4]	Files	Real
Franklin et al. [31]	Attack patterns	Real
Romero-Gomez et al. [66]	DNS records, WHOIS records, malware, domain blacklist	Real
Syamkumar et al. [78]	Border Gateway Protocol updates	Real
Assal et al. [12]	Source code	Real
Arendt et al. [10]	Network services logs	Real
Siadati et al. [73]	Login events	Real
Alam et al. [2]	Program	Real
Buchanan et al. [17]	N/A	N/A
Anh Huynh et al. [9]	Network traffic	Real and simulated
Capper and van Wijk [20]	Network traffic	Real and simulated
Gove [33]	Security policies	Real
Aupetit et al. [13]	Network traffic	Real
Peryt et al. [62]	Top-level-domain data	Real
Post et al. [63]	Software-defined network data	Simulated

Table 3: 60 VizSec papers classified according to the security task they aim to achieve

Forensic Analysis	[18], [16], [37], [84], [19], [46], [83], [35], [81], [49], [78], [73], [20], [13]
Threat Analysis	[57], [69], [48], [30], [6], [31], [66]
Malware Analysis	[60], [7], [68], [4], [9], [62]
Security Awareness and Education	[21], [61], [59], [58]
Privacy Awareness	[41], [70], [26], [23]
Situational Awareness	[10], [63]
Vulnerability Management	[5], [64], [27], [3], [12], [2]
User Behaviour Analytics	[36], [77], [11], [87], [22]
Incident Handling	[34], [14], [17]
Security Management	[8], [33]
Triage Analysis	[51]
Firewall Rules Analysis	[44]
Methodology	[85], [71]
Machine Learning	[74], [15], [75]

The most common security task that visualizations help with is the process of forensic analysis. *Forensic Analysis* involves investigating incidents that pose a threat to the organisation to gain a better understanding of the perpetrators and their capabilities [53]. Among forensic analysis, network forensics is the most widely used. It deals with the analysis of network activities to determine the source of security policy violations or information security breaches [55]. In addition to forensic analysis, there is also *Triage analysis*, which is the most basic phase in the analysis process. It includes the tasks of eliminating noise in the raw data, identifying and grouping the data that indicate suspicious events worthy of further investigation [90].

Visualization can also be very useful in *Threat Analysis*, which involves activities that help identify, analyze and prioritise potential security threats to a system and the information processed within it [82]. Much of the work deals with specific threats known as malware. More specifically, with the analysis of samples and various files that are potentially malicious, which we call *Malware Analysis*. In addition to malware, some of the work also deals with *Vulnerability Management*, whether to identify vulnerabilities or to support decision making. An important part of security is *Incident Handling*, which shows us how organisations or individuals respond to an attack. In this context, visualization can help with decision-making and identifying potential threats to the system under observation.

Visualization can also be useful in analyzing user behaviour and activity, which is known as *User Behaviour Analytics* and is often associated with insider threat detection. Work that deals with the implementation and understanding of various security policies and specifications is classified as *Security Management*.

A very important function of visualization is also to raise awareness in the field of cyber security. Here we can distinguish three areas of awareness. The first is the most general, which includes various demonstrations and Capture the Flag (CTF) games, not only to raise awareness but also to fulfil an educational function, we call it *Security Awareness and Education*. The second is *Situational Awareness*, which aims to provide insight into and raise awareness of the state of a system or network. And finally *Privacy Awareness*, which deals with sensitive private data and tries to raise awareness and reduce the risk of its disclosure through visualizations.

As a separate category, we have highlighted the work that deals with the analysis of firewall rules, hence the category is called *Firewall Rules Analysis*. Of the other works, we have singled out two further groups. The first includes papers that deal with visualization methodology and models for effective visualization, in short we call this group *Methodology*. The second group deals exclusively with *Machine Learning*, i.e. a better interpretation of the model used.

All 60 VizSec papers were categorized according to the security tasks described and presented in the Table 3. From this table, we can conclude that in recent years, when it comes to contributions to VizSec, visualization is most often used in forensic analysis tasks, in about a quarter of all papers, especially in network forensics. The reason for this is that hidden patterns in network traffic data are easier for the analyst to detect with appropriate visualization than when viewing raw data. Network traffic is mostly used as input data for this security task.

The second major area where visualization is used is threat and malware analysis, with just over 20% of work falling into this category. Here, malware samples, attack patterns, and files (source codes and executables) are most commonly used. In most cases some attributes of the mentioned input data are visualised to further improve their analysis.

The third main group is awareness and education, with the categories of general security awareness, privacy awareness, and situational awareness accounting for about 16% of the contributions. These tasks lend themselves very well to visualization, as it is a great tool to demonstrate some security aspects to users who are not security experts [21, 61] or to visualize CTF game events for additional feedback [59]. Privacy awareness is supported through visualization of collected personal data [70] and decision support for data sharing using differential privacy [41]. It also discusses how to minimize disclosure risks in visualizations [26] and how to perceptually mask privacy in graphs [23]. Increasing situational awareness is achieved by visualizing the entire system or network at a high level [10, 63]. The input data for these tasks is often some personal or sensitive data, in addition to network traffic data and machine learning datasets.

As for vulnerability management, six papers deal with this task. Here, visualization is used to identify vulnerabilities [2], eliminate vulnerabilities in code [12], and improve the decision-making process in assessing them [3]. Vulnerabilities, source codes and binary applications are mostly used for these tasks.

In five papers, user behavior analytics is the main security task. These works emphasize insider threat detection as the main goal [11, 36], mainly using user behavior logs and events. Here, visualization is used to explore and identify user behavior patterns and understand why some behaviors are considered anomalous [22], and to build mental models about user activities [11].

For incident handling, visualization is useful by providing the dashboard with various views around the system, its requirements, and the associated attack vector space [14]. In addition, Buchanan et al. [17] described what work in the incident handling process could benefit from visualization. Visualization is also used to generate compact representations of cyber narratives with the goal of helping analysts navigate the relationships between key victims and attackers [34]. The data used for this task consists of attack patterns, vulnerabilities, weaknesses, and incident reports.

When it comes to security management, visualization helps in adopting frameworks [8] or policies [33], described in two papers that use specifications and security policies as inputs.

Only one work was found that supports the task of triage analysis, and one that supports firewall rules analysis. The first work focuses on providing a dashboard to display OT network sensor values to provide insight into the data and support triage analysis [51]. The second work uses visualization to automatically analyze the current control conditions of packets and displays the conditions so they can be easily verified [44].

We have placed papers that do not have a security component built into their input data, but provide a specific methodology in an area of visualization in cybersecurity, in the methodology category. Varga et al. [85] aimed to develop a methodology for the development of a military cyber symbology to enable the visualization of cyber situations. On the other hand, Sethi and Wills [71] dealt with the creation of a model for effective visualization in cybersecurity. Of course, these two works do not have specific input data as a dataset, but use some other researches for their work.

Similar to the methodological work, the work we have categorized as machine learning does not have a specific security task in mind,

Table 4: 55 VizSec papers classified based on geometry used

2D without CMV	[34], [57], [64], [3], [21], [60], [59], [87], [58], [66], [78], [12], [73], [2], [33]
2D with CMV	[74], [5], [36], [69], [18], [27], [70], [16], [15], [61], [37], [48], [77], [84], [30], [51], [7], [75], [11], [19], [14], [6], [22], [83], [35], [8], [68], [81], [4], [31], [10], [9], [20], [13], [63]
3D without CMV	[44], [49], [62]
3D with CMV	[41], [46]

but focuses only on improving the machine learning component. Visualization in this area helps developers of deep learning models better understand how to interpret their models and better identify misclassifications [15]. In addition, visualization helps with decision making [75] and provides an overview of how machine learning models work and how data is evaluated over time as input data is added [74].

4.3 Visual Encoding

The most important component of any visualization system is the visual encoding itself, that is, the visualization techniques used to achieve the desired representation.

Most of the 60 selected VizSec papers focus on the development of visualization systems that support a specific security task. Five papers are not about the creation of visualizations, but discuss problems, shortcomings, and standards of some visualizations. These papers address military cybersecurity [17, 85], privacy in visualization [23, 26], and description of a model for effective visualization in cybersecurity [71].

Based on this, in the remainder of this section, we have classified the works that focus on creating visualization systems based on the geometry and visualization techniques used, as well as the visualization categories into which these techniques belongs to.

4.3.1 Geometry

First, we categorized visual encoding at a very high level. We divided the works into three broad classes based primarily on the geometry used. 2D visualizations are the simplest and are represented by some graphs or diagrams that stand alone. If we extend the 2D visualization with multiple graphs connected by interactions, we get a slightly more complicated view. Therefore, we have classified such visualizations into the 2D visualizations with coordinated multiple views (CMV) technique. The main feature of the CMV technique is to provide the user interactions with information and different representations to better understand the displayed data [65]. The most complicated visualization that appears is, of course, the three-dimensional one and forms its own group.

Table 4 shows the classification of papers from VizSec, which deal with the creation of visualization systems, into the categories based on the used geometry. The table shows that three-dimensional representations are rarely used in cybersecurity visualization, only in five papers. For two-dimensional representations, the CMV technique is more prevalent than simple 2D visualizations because it creates a connection between different aspects of the data, which ultimately gives the user a greater opportunity to interact with and gain insights from the data. Of the selected papers that use a 2D geometry representation, 64% use the CMV technique and 36% do not, which is a high percentage considering its advantages and popularity. Interestingly, among the works that use 3D visualization techniques, Krokos et al. [46] and John et al. [41] also used the CMV technique.

4.3.2 Visualization Techniques

After looking at the visual encoding from the perspective of the geometry used, we analyzed the specific visualization techniques that were used. In addition to the visual techniques, we also observed the groups of these techniques. The groups and the names of the techniques themselves are inspired by the website d3 graph gallery [39]. This website is based on visualization techniques available in the d3.js library, which is most commonly used for creating information visualizations. Therefore, most of the visualizations in the analyzed works can be directly related to the groups and techniques listed there.

However, not all techniques that appear in the VizSec papers are represented on the aforementioned website, which may be due to the fact that some of them are completely different from those listed there or consist of a combination of several techniques. Therefore, we have made some modifications and assumptions before showing the results of the analysis.

The Spiral Plot [51] and Sunburst [7, 87] techniques are similar to the Circular Barplot and Doughnut techniques. We have categorized the Sunburst technique under Circular Barplot. On the other hand, we had to put the Spiral Plot technique in the Other category because of its characteristics. The Icicle Plot technique [19, 69] is a combination of the Dendrogram technique and the Treemap technique and is also categorized under Other. Flow techniques were used by Krokos et al. [46], Chen et al. [22], and Subramanian et al. [77]. These techniques are much more complex than anything offered under the group of the same name, so they have also been categorized under Other. Similar to the flow techniques, Gove [34] has created a visualization based on the Gantt chart. Chaffey and Sgandurra [21] used the concept of a 2D game for visualization, which is also categorized under Other.

We have identified a variant of the Radar chart that uses points [4] but is classified as a Radar chart. The same is true for the Flatten Doughnut chart [10], which looks like a simple bar and is classified as a Doughnut chart. Theron et al. [81] used a visualization technique called Hive Plot, but on closer inspection it is only a 3D extension of the Arc diagram and is therefore classified as such.

In addition to grouping some diagrams in the Other class, we have added some new classes to classify as many visualization techniques as possible. Gove [33], Angelini et al. [4] and Reynolds et al. [64] used a visualization similar to the Heatmap, but with points and no correlation element. These charts are often referred to as Dot matrix charts, a term we have also used. Arendt et al. [11] focused on the glyph visualization technique and therefore is covered in a separate class. Various diagram-like visualizations that resemble control flow diagram are also present [2, 7, 20, 57, 68]. These techniques are mostly a fusion of dendrograms and network techniques, but cannot be classified as either technique due to their hierarchical and fluid nature. Therefore, these visualizations are placed in a new group called Diagram.

In this categorization, we did not distinguish between 2D and 3D visualizations, but assigned the 3D visualizations to the corresponding 2D techniques. For each visualization technique indicated on the d3 graph gallery website, we calculated the percentage of occurrence in selected VizSec papers. The calculated percentages are shown in Fig. 1.

From Fig. 1 it is easy to see that Barplot is the predominant technique among the selected works, with over 35% occurrences. This is followed by the techniques Histogram and Heatmap, each with around 25% of the uses. The techniques Scatter and Network are also represented with over 20% of uses each. Other techniques are present in about 15% of papers and Connected scatter in around 10%. All other 20 discovered visualization techniques appear in the observed VizSec papers with less than 10% each.

4.3.3 Visualization Categories

Having analyzed the techniques themselves, we now turn to the analysis of the visualization categories. These categories were also created based on the d3 graph gallery website. Again, we made some changes to classify as many techniques as possible. All techniques described as variations of a particular technique are categorized as that primary technique. Also, we have divided the techniques that

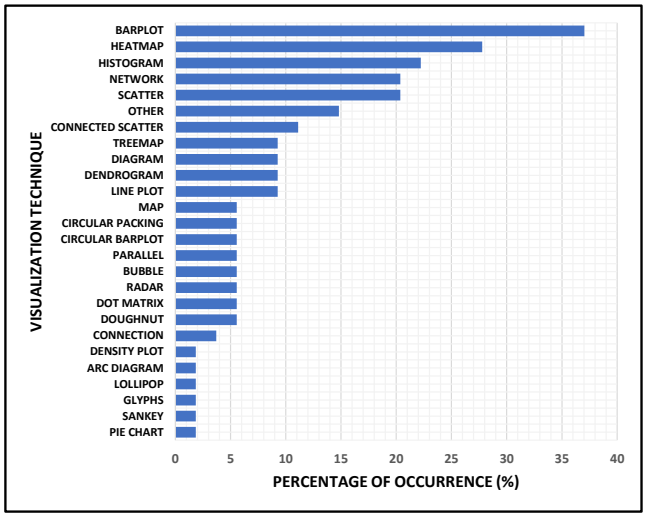

Figure 1: Percentage of occurrence of each visualization technique in the 55 VizSec papers

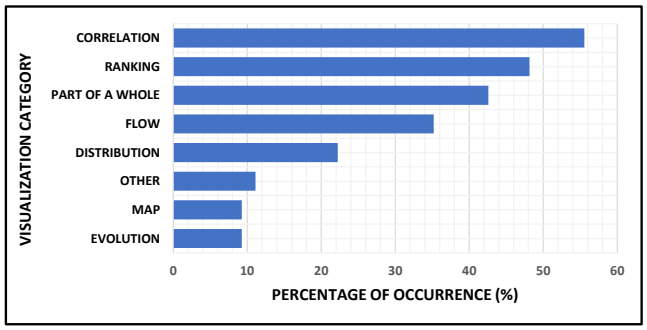

Figure 2: Percentage of occurrence of each visualization category in the 55 VizSec papers

resulted from the fusion of several techniques into several categories based on the primary techniques that compose them.

Fig. 2 shows a bar chart of the proportion of visualization categories in VizSec papers that create visualization systems in their work. The diagram shows that the Correlation techniques are predominant with over 55% of uses in the papers, while Ranking and Part of a whole techniques have about 45% of occurrence in these papers each. Next to them are Flow techniques with a share of about 30% and Distribution techniques with 22%. At the bottom of the diagram are Evolution, Map, and techniques classified as Other.

4.4 Interactivity

Interaction is one of the most important components of any modern visual system. To ensure that visualization is not limited to displaying data, we give the user the ability to manipulate the display and data through interaction. Interaction serves as a dialog between the user and the system as the user explores the information presented. When attempting to categorize interactions, the problem arises that the proposed taxonomies vary in their granularity.

Since there are numerous interaction techniques in each work and many of them do not deal with the interactions of the created visual system, it is a difficult task to classify each work based on the interaction techniques. To overcome this problem, the VizSec papers were categorized based on the goal of interaction in the context of the whole system and visual analytics.

Visual analytics uses visualization and interaction techniques to integrate human judgment into the data analysis process [25]. In this context, we have divided systems that use interactions into exploratory and expressive systems [28].

Table 5: 60 VizSec papers classified by visual analytics type

No	[34], [57], [21], [61], [85], [26], [60], [23], [87], [71], [78], [17], [62]
Exploratory	[36], [69], [64], [18], [27], [70], [16], [3], [15], [77], [84], [30], [51], [59], [19], [14], [6], [22], [46], [83], [35], [8], [44], [68], [49], [31], [66], [12], [10], [73], [2], [33], [13], [63]
Expressive	[74], [5], [41], [37], [48], [7], [75], [11], [58], [81], [4], [9], [20]

Exploratory systems focus on exploring the space of the visualization itself and the information displayed within it. These include zooming, displaying details of a set of information, filtering, highlighting, grouping, etc. Such interactions give the user insight into the data by performing interactive actions with it and observing its behaviour. Exploratory interactions are closely related to the CMV visualization technique, as it is based on this type of interaction [65].

Expressive systems rely on interaction to completely change the algorithm on which the visualization or data analysis is based, whether by completely changing the visualization representation or just changing the model on which the data analysis is based [25]. These systems often work according to the "human in the loop" principle, i.e. the user gains insights by using the visualization as a tool and interacting with it to manipulate and improve the existing model [67]. These types of systems are almost always based on a machine learning algorithm. Some examples of interactions in these systems are manually revising the model predictions to create a better model, or changing the parameters of the model that are directly responsible for generating the data.

In addition to these two systems, we have distinguished which works do not use interactions at all or, as such, do not fall within the scope of visual analytics. Based on this categorization, we created the Table 5, where the VizSec papers are arranged based on the classification described.

The table shows that 13 works do not belong to the field of visual analytics, that is, their visualization system does not have any type of interaction behind it, giving a percentage of 78% of the works with some type of interaction. Of the works that fall into the visual analytics domain, we can classify 72% as exploratory types and 28% as expressive types.

4.5 Evaluation

Once the visualization system is created, it must be verified in some way for accuracy, usefulness, and quality. This is precisely the goal of the last step of the methodology for creating cybersecurity visualizations we consider in this paper: evaluation. Evaluation can be done during the development of the system or after the process has been completed. Formative evaluation takes place during the development of the visualization system for the purpose of improvement in order to guide the further development process in the right direction. Summative evaluation, on the other hand, is performed after a system has been completed to identify deficiencies and errors that can be corrected in the future or recorded for future projects.

In this paper we have looked at which papers have an evaluation, whether formative or summative, and how this has been achieved. Of course, we cannot expect every evaluation method to be the same, but we can pick out a limited number of generalized evaluation methods.

The most commonly used evaluation methods are the use case and the case study. These two seemingly similar concepts differ in that the use case examines the potential application of a developed system, while the case study examines an actual application of the system in the real world. Interviews, surveys, and feedback, whether from experts or non-experts are also used as a form of evaluation.

Table 6: 60 VizSec papers classified by methods of formative and summative evaluation

Paper	Formative evaluation	Summative evaluation
Sopan and Berlin [74]	Non-professional feedback	Use case
Gove [34]	Professional feedback	Professional feedback, System testing and analysis
Angelini et al. [5]	Professional interview	Case study
John et al. [41]	✗	Professional study
Graham et al. [36]	Professional interview	Case study
Nadeem et al. [57]	✗	System testing and analysis, Use case
Schreiber et al. [69]	✗	Case study, Use case
Reynolds et al. [64]	Professional study and feedback	Use case, Professional interview
Böhm et al. [18]	Professional interview	Use case
Dennig et al. [27]	Professional interview	Use case, Professional feedback
Schufrin et al. [70]	✗	Use case, Non-professional study
Beran et al. [16]	Professional interview	Professional study
Alperin et al. [3]	✗	Use case
Becker et al. [15]	Professional interview	System testing and analysis
Chaffey and Sgandurra [21]	✗	✗
Peng et al. [61]	Non-professional feedback	✗
Guerra et al. [37]	Professional interview	Non-professional study
Laughlin et al. [48]	✗	Use case
Varga et al. [85]	✗	✗
Subramanian et al. [77]	✗	Case study
Dasgupta et al. [26]	✗	Case study
O'Shaughnessy [60]	✗	System testing and analysis
Ulmer et al. [84]	✗	Use case, Professional feedback, Non-professional study
Fouss et al. [30]	✗	Case study, Professional feedback
Lohfink et al. [51]	Professional interview	Use case, Professional interview, Non-professional study
Angelini et al. [7]	Professional feedback	Use case
Oslejšek et al. [59]	Professional interview	Non-professional study
Chou et al. [23]	✗	Non-professional study
Sopan et al. [75]	Professional interview	Professional feedback
Arendt et al. [11]	Professional interview	Non-professional study
Cappers et al. [19]	✗	Use case
Bakirtzis et al. [14]	✗	Use case
Angelini et al. [6]	Non-professional study	Case study
Yang et al. [87]	✗	Professional study
Chen et al. [22]	Professional feedback	Case study
Krokos et al. [46]	Professional interview	Professional feedback
Ulmer et al. [83]	Non-professional study, Professional feedback	Use case
Gove and Deason [35]	Professional interview and feedback	Use case
Norton and Qi [58]	✗	System testing and analysis
Angelini et al. [8]	Professional feedback	✗
Sethi and Wills [71]	Professional interview	✗
Kim et al. [44]	✗	Case study, Professional interview
Santhanam et al. [68]	✗	Case study
Theron et al. [81]	✗	Case study
Leichtnam et al. [49]	✗	Case study
Angelini et al. [4]	✗	Case study
Franklin et al. [31]	Professional interview	✗
Romero-Gomez et al. [66]	✗	Use case, Professional study and interview
Syamkumar et al. [78]	✗	Use case
Assal et al. [12]	Professional study	Non-professional study
Arendt et al. [10]	✗	Non-professional feedback, Case study, Professional interview
Siadati et al. [73]	✗	Professional study
Alam et al. [2]	✗	System testing and analysis
Buchanan et al. [17]	Professional survey	Professional interview
Anh Huynh et al. [9]	✗	Use case
Capper and van Wijk [20]	✗	Use case
Gove [33]	Professional survey	Use case
Aupetit et al. [13]	✗	Use case
Peryt et al. [62]	✗	✗
Post et al. [63]	✗	✗

Such methods are most often used in formative evaluations where we want to get quick help and advice on the current state of the system in order to make timely corrections.

Methods such as user studies are time consuming and require significant human resources, but provide a complete evaluation response in terms of accuracy, usefulness, and quality of the system. Therefore, they are often conducted as part of a summative evaluation. There are several papers that have done evaluation using test cases, performance tests or model analysis. Although these are not true evaluation methods, they are presented here as a form of system testing and analysis.

Based on the participants taking part in the evaluation, we have considered two main groups. The first group is referred to as non-professionals, i.e. users who have no experience in a particular field, such as students, the general population or people with a limited technical background only. In contrast, there are professionals who have been working in a particular field for a long time and are experienced.

Based on these selected evaluation methods, the Table 6 was created, which contains a list of the 60 VizSec papers and the evaluation method used (or not used) in them, in the form of a formative and summative evaluation. We can see that 55% of the papers have no formative evaluation, while this percentage is 13% for summative

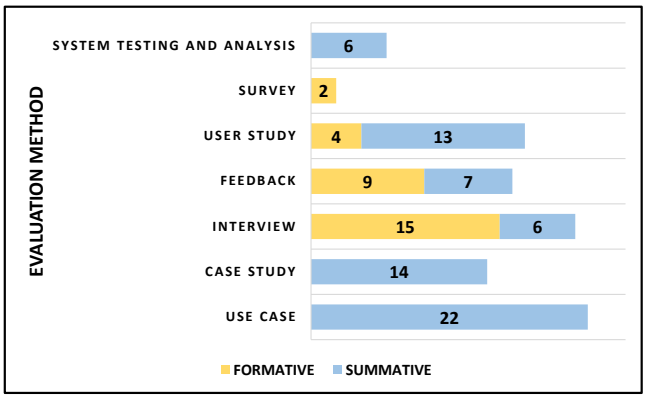

Figure 3: Number of papers classified in each of generalized evaluation method in the 60 VizSec papers

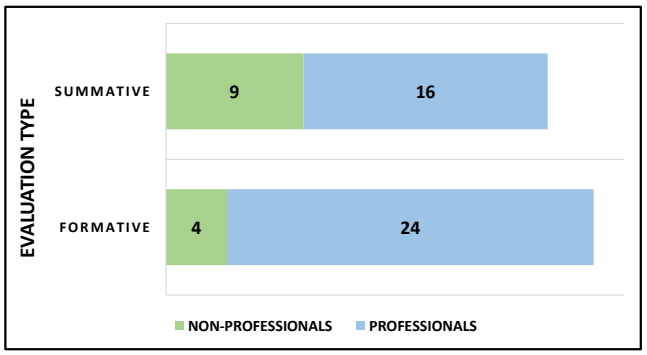

Figure 4: Number of papers classified by the type of participants involved in the evaluation in the 60 VizSec papers

evaluation. Only four papers contain no evaluation at all, while 23 papers used both formative and summative evaluation.

The Fig. 3 shows the number of papers with each generalised evaluation method in the context of formative and summative evaluation. The chart shows that case studies and use cases play a predominant role in summative evaluation, with user studies right next to them. Both case studies and use cases only occur in summative evaluation, as do system tests and analyzes. In formative evaluation, interview and feedback methods have the greatest impact, as opposed to user studies and surveys. The survey method also does not occur in summative evaluation.

The Fig. 4 shows the number of papers with formative and summative evaluation based on the two main groups of participants in the evaluation. The chart shows that professionals are the main group for both formative and summative evaluation, with a higher percentage in formative evaluation.

Looking at both the evaluation method and the type of participants involved in the evaluation, the most commonly used formative evaluation is the professional interview, which is used in over 50% of the cases where some type of formative evaluation is used. This is followed by professional feedback with just under 26% and non-professional feedback with about 7% of usage. In terms of summative evaluation, use case is the most common with 42% of uses, followed by case study with about 27%. In addition, the non-professional study is the only one that accounts for more than 15% of the uses in the summative evaluation.

5 Discussion

In the previous section, we categorized the VizSec papers based on the selected attributes of each component of the visualization design. Now we will highlight the main results of the analysis and present some interesting patterns and correlations that emerged from the relationships between the observed categories.

In analyzing the input data, we found that most of the work is based on a real data source. When this is not the case, it is often sensitive data that is used. The papers that are based on a simulated data source are mainly concerned with privacy awareness and user behaviour analysis. In the selected VizSec papers the input dataset consists mainly of network traffic and logs. In addition, malware samples and files in the form of executables or source files are also frequently used.

Of the 14 security tasks identified, visualization was mainly used in forensic analysis. It is also commonly used in threat and malware analysis, as well as for awareness and education purposes. In contrast, incident handling and security management are two tasks that are not as commonly supported by visualization, according to our research. Of the selected VizSec papers, those that do not focus on creating visualizations mostly relate to methodology and security management tasks.

As this study shows, hardly any of the selected visualization solutions are implemented in 3D. Adding another dimension when creating a visual solution can be beneficial for certain tasks, but must be used wisely to ensure that the final visualization provides enough detail without being too cluttered and difficult to use. When it comes to works that use 2D visual solutions, more than 60% of the selected works also uses the CMV technique. It is certainly a good trend to offer the user more different visual representations associated with interactions. Likewise, it is a good practice to include interactions that allow users to make dynamic changes to the visual representation as they use it. In this sense, it is certainly encouraging to see that more than three-quarters of the works include some kind of interaction. Of these, more than 70% are classified as exploratory types and the rest as expressive types.

Many visualization techniques that appear in the selected papers are similar and come in the form of bar charts, histograms, heat maps, networks, and scatter techniques. In contrast, some techniques such as pie charts, chord diagrams, maps, etc. are rarely or never used. In this sense, it is encouraging to use some of these rarely used techniques or to develop new ones.

This research has shown that almost all of the papers contain some kind of evaluation. In most cases, these are summative evaluations, which are included in almost 90% of the selected papers, while less than half of all papers contain formative evaluations. With this result, we encourage researchers to use formative evaluations more frequently so that they can more easily identify the most important problems and challenges that need to be addressed and eventually develop visual solutions that can be more effective for end users with the help of experts from the target field.

Our study showed that case studies and use cases are mainly used in summative evaluation. Therefore, we recommend using more user studies in this type of evaluation as they can provide additional feedback on the usage of the created visualization solution. For formative evaluation, mainly interview and feedback methods are used. In both types of evaluation, experts play a larger role than non-experts. In summative evaluation, it can sometimes be advantageous to involve a large group of non-experts in the evaluation process, as they can provide different feedback than the experts, especially if the final product is not intended only for professional use. Interestingly, our research suggests that papers with formative evaluation are the most likely to use the CMV technique, while the others do not. It is also interesting to note that papers with a higher level of visual analytics also contain a higher level of evaluation.

All of these conclusions were drawn for only the last six years of VizSec contributions, so these results should not be generalized to the entire field of cybersecurity visualization. Instead, these results should provide an indication of the direction in which visualization is trending in the field, i.e., which parts of the field are rarely explored and are open to further scientific research.

6 Related work

As mentioned in Sect. 3, we were able to find 17 surveys that dealt in some way with visualization in cybersecurity. Most of these survey papers addressed only two or three identified focal points in the creation of cybersecurity visualizations, i.e., components of visualization design. These identified components are closely related to those in other works [56] and are not specific to the target domain, but are general to any visualization solution. To our knowledge, this is a novel idea to collect all survey papers in the field of cybersecurity visualization and point out which of these papers are focused on.

Among the survey papers that cover most of the components of visualization design, Shiravi et al. [72] and Zhang et al. [88] provided an overview of visualization in cybersecurity from the aspects of Input Data, Visual Encoding, Use Cases and Applications, and Tasks and Challenges. In contrast, Kartel et al. [42] and Wagner et al. [86] focused on the same aspects, but in addition to Interactivity rather than Use Cases and Applications. Finally, Liu and He [50] again focused on Input Data and Visual Encoding, but also on Use Cases and Applications as well as Evaluation. In our research, we focused on the components of Input Data, Visual Encoding, Interactivity and Evaluation. In addition, we included the Security Tasks component, which can be considered as a smaller part of the Use Cases and Applications and Tasks and Challenges components.

In classifying the visualization solutions created as part of the VizSec contributions, we used some novel ideas to classify the individual contribution within each component of the visualization design. For Input Data and Security Task component we did not use a strict categorization. For the Input Data, we created a list of data types that appear in the works, along with the attribute of data source fidelity. As for the Security Tasks, when analyzing the VizSec works, we identified 14 categories based on the security tasks supported by the visualization. On the other hand, for the Visual Encoding, we used a new type of categorization based on the visual techniques and categories of the d3 gallery. For the interactivity component, we also used a new type of categorization based on the level and type of visual analysis, inspired by the classification of Endert et al. [28]. For the assessment component, we showed how the assessment was or was not achieved for each VizSec paper in terms of formative and summative assessment.

Regarding work that focuses on VizSec papers from multiple years, three existing papers focus only on VizSec papers. Staheli et al. [76] classified ten years of VizSec papers by dimensions and evaluated components of the visualization system and techniques of evaluation. Crouser et al. [24] identified common thematic groupings with a similar number of selected papers based on cosine similarity between the TF-IDF vectors constructed for each paper. Goodall [32] focused only on VizSec articles whose visualization supports the network defense theme. He analyzes possible data sources for network defense as well as the visualization solutions themselves. It also discusses all VizSec 2017 articles in the context of the topics they contain.

7 Conclusion

In this paper, we found and analyzed 17 survey papers on the topic of visualization in cybersecurity. The papers were analyzed to find out what they focus on when it comes to aspects of creating cybersecurity visualizations. Based on this, we identified six categories as areas of focus for these survey papers. For each of these categories, we indicated whether or not it was addressed in the survey. Based on the identified categories, we derived five components of the

visualization design. These components include: the data used as input to the visualization, the task in the security domain to be solved by the visualization, the visual encoding chosen, the degree of interactivity used to accomplish the target task, and the evaluation of the visualization system created. Based on the five selected components, we categorized 60 papers from the VizSec Symposium from 2016 to 2021 to perform an analysis of the design space of the visualization systems presented in these papers.

In addition to identifying the type of input data used in VizSec papers, we also categorized the data sources used based on the fidelity attribute. In terms of security tasks, we identified 14 different tasks where visualization can be used and categorized the papers accordingly. Regarding visual encoding, we categorized contributions by geometry, visualization categories, and visualization techniques used. To classify papers based on interactions, we observed the use of visual analytics and its type in the selected papers. In the final analysis, the evaluation methods presented in the selected papers were generalized and classified into different categories based on the type of participants involved and the type of evaluation, for both formative and summative evaluation.

As our analysis of surveys in the Sect. 3 shows, the interactivity has been little studied. In the future, it may be beneficial to closely analyze papers in terms of interactivity, not only to examine low-level interaction techniques, but also to categorize them and show which categories are most commonly used.

Based on the results of this research, the next step could be to develop a methodology for creating cybersecurity visualizations. For each security task, we could create a set of guidelines and provide a set of use cases and applications to show real-world examples. Thus, when creating the visualization, one can use the guidelines created to make the visualization more effective and standardized in that security domain, or to verify that a visualization created meets the guidelines and use that knowledge to evaluate the validity of the visualization.

Acknowledgments

This work has been supported by the European Union's European Regional Development Fund, Operational Programme Competitiveness and Cohesion 2014-2020 for Croatia, through the project Center of competencies for cyber-security of control systems (CEKOM SUS), grant KK.01.2.2.03.0019.

References

[1] C. N. Adams and D. H. Snider. Effective data visualization in cybersecurity. In *SoutheastCon 2018*, pp. 1–8, 2018. doi: 10.1109/SECON.2018.8479113

[2] M. J. Alam, M. T. Goodrich, and T. Johnson. J-viz: Finding algorithmic complexity attacks via graph visualization of java bytecode. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739575

[3] K. B. Alperin, A. B. Wollaber, and S. R. Gomez. Improving interpretability for cyber vulnerability assessment using focus and context visualizations. In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 30–39, 2020. doi: 10.1109/VizSec51108.2020.00011

[4] M. Angelini, L. Aniello, S. Lenti, G. Santucci, and D. Ucci. The goods, the bads and the uglies: Supporting decisions in malware detection through visual analytics. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062199

[5] M. Angelini, G. Blasilli, S. Bonomi, S. Lenti, A. Palleschi, G. Santucci, and E. D. Paoli. Bucephalus: a business centric cybersecurity platform for proactive analysis using visual analytics. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 15–25, 2021. doi: 10.1109/VizSec53666.2021.00007

[6] M. Angelini, G. Blasilli, P. Borrello, E. Coppa, D. C. D'Elia, S. Ferracci, S. Lenti, and G. Santucci. Ropmate: Visually assisting the creation of rop-based exploits. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709204

[7] M. Angelini, G. Blasilli, L. Borzacchiello, E. Coppa, D. C. D'Elia, C. Demetrescu, S. Lenti, S. Nicchi, and G. Santucci. Symnav: Visually assisting symbolic execution. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–11, 2019. doi: 10.1109/VizSec48167.2019.9161524

[8] M. Angelini, S. Lenti, and G. Santucci. Crumbs: A cyber security framework browser. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062194

[9] N. Anh Huynh, W. Keong Ng, A. Ulmer, and J. Kohlhammer. Uncovering periodic network signals of cyber attacks. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739581

[10] D. Arendt, D. Best, R. Burtner, and C. Lyn Paul. Cyberpetri at cdx 2016: Real-time network situation awareness. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2016. doi: 10.1109/VIZSEC.2016.7739584

[11] D. L. Arendt, L. R. Franklin, F. Yang, B. R. Brisbois, and R. R. LaMothe. Crush your data with vic2es then chissl away. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709212

[12] H. Assal, S. Chiasson, and R. Biddle. Cesar: Visual representation of source code vulnerabilities. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739576

[13] M. Aupetit, Y. Zhauniarovich, G. Vasiliadis, M. Dacier, and Y. Boshmaf. Visualization of actionable knowledge to mitigate drdos attacks. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739577

[14] G. Bakirtzis, B. J. Simon, C. H. Fleming, and C. R. Elks. Looking for a black cat in a dark room: Security visualization for cyber-physical system design and analysis. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709187

[15] F. Becker, A. Drichel, C. Müller, and T. Ertl. Interpretable visualizations of deep neural networks for domain generation algorithm detection. In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 25–29, 2020. doi: 10.1109/VizSec51108.2020.00010

[16] M. Beran, F. Hrdina, D. Kouřil, R. Ošlejšek, and K. Zákopčanová. Exploratory analysis of file system metadata for rapid investigation of security incidents. In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 11–20, 2020. doi: 10.1109/VizSec51108.2020.00008

[17] L. Buchanan, A. D'Amico, and D. Kirkpatrick. Mixed method approach to identify analytic questions to be visualized for military cyber incident handlers. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739578

[18] F. Böhm, L. Englbrecht, S. Friedl, and G. Pernul. Visual decision-support for live digital forensics. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 58–67, 2021. doi: 10.1109/VizSec53666.2021.00012

[19] B. C. Cappers, P. N. Meessen, S. Etalle, and J. J. van Wijk. Eventpad: Rapid malware analysis and reverse engineering using visual analytics. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709230

[20] B. C. M. Cappers and J. J. van Wijk. Understanding the context of network traffic alerts. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739579

[21] E. J. Chaffey and D. Sgandurra. Malware vs anti-malware battle - gotta evade 'em all! In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 40–44, 2020. doi: 10.1109/VizSec51108.2020.00012

[22] S. Chen, S. Chen, N. Andrienko, G. Andrienko, P. H. Nguyen, C. Turkay, O. Thonnard, and X. Yuan. User behavior map: Visual exploration for cyber security session data. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2018. doi: 10.1109/VIZSEC.2018.8709223

[23] J.-K. Chou, C. Bryan, J. Li, and K.-L. Ma. An empirical study on

perceptually masking privacy in graph visualizations. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709181

[24] R. J. Crouser, E. Fukuda, and S. Sridhar. Retrospective on a decade of research in visualization for cybersecurity. In *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–5, 2017. doi: 10.1109/THS.2017.7943494

[25] W. Cui. Visual analytics: A comprehensive overview. *IEEE Access*, 7:81555–81573, 2019. doi: 10.1109/ACCESS.2019.2923736

[26] A. Dasgupta, R. Kosara, and M. Chen. Guess me if you can: A visual uncertainty model for transparent evaluation of disclosure risks in privacy-preserving data visualization. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161608

[27] F. L. Dennig, E. Cakmak, H. Plate, and D. A. Keim. Vulnex: Exploring open-source software vulnerabilities in large development organizations to understand risk exposure. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 79–83, 2021. doi: 10.1109/VizSec53666.2021.00014

[28] A. Endert, C. Han, D. Maiti, L. House, S. Leman, and C. North. Observation-level interaction with statistical models for visual analytics. In *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 121–130, 2011. doi: 10.1109/VAST.2011.6102449

[29] D. Ferebee and D. Dasgupta. Security visualization survey. In *Proceedings of the 12th Colloquium for Information Systems Security Education University of Texas*, p. 124. Citeseer, 2008.

[30] B. Fouss, D. M. Ross, A. B. Wollaber, and S. R. Gomez. Punyvis: A visual analytics approach for identifying homograph phishing attacks. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161590

[31] L. Franklin, M. Pirrung, L. Blaha, M. Dowling, and M. Feng. Toward a visualization-supported workflow for cyber alert management using threat models and human-centered design. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062200

[32] J. R. Goodall. *Introduction to Visualization for Computer Security*, pp. 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-78243-8_1

[33] R. Gove. V3spa: A visual analysis, exploration, and diffing tool for selinux and seandroid security policies. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739580

[34] R. Gove. Automatic narrative summarization for visualizing cyber security logs and incident reports. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–9, 2021. doi: 10.1109/VizSec53666.2021.00005

[35] R. Gove and L. Deason. Visualizing automatically detected periodic network activity. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709177

[36] M. Graham, R. Kukla, O. Mandrychenko, D. Hart, and J. Kennedy. Developing visualisations to enhance an insider threat product: A case study. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 47–57, 2021. doi: 10.1109/VizSec53666.2021.00011

[37] J. L. Guerra, E. Veas, and C. A. Catania. A study on labeling network hostile behavior with intelligent interactive tools. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161489

[38] T. Haina, H. Chunjing, and G. Jingguo. Applications of visualization technology for network security. In *2017 IEEE Trustcom/BigDataSE/ICESS*, pp. 1038–1042, 2017. doi: 10.1109/Trustcom/BigDataSE/ICESS.2017.349

[39] Y. Holtz. The D3 Graph Gallery. `https://d3-graph-gallery.com/`, 2018. Accessed: 20.6.2022.

[40] S.-Y. Ji, B.-K. Jeong, and D. H. Jeong. Evaluating visualization approaches to detect abnormal activities in network traffic data. *International Journal of Information Security*, 20(3):331–345, 2021.

[41] M. F. S. John, G. Denker, P. Laud, K. Martiny, A. Pankova, and D. Pavlovic. Decision support for sharing data using differential privacy. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 26–35, 2021. doi: 10.1109/VizSec53666.2021.00008

[42] A. Kartel, E. Novikova, and A. Volosiuk. Analysis of visualization techniques for malware detection. In *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 337–340. IEEE, 2020.

[43] R. R. Kasemsri. A survey, taxonomy, and analysis of network security visualization techniques. Master's thesis, Georgia State University, 2006.

[44] H. Kim, S. Ko, D. S. Kim, and H. K. Kim. Firewall ruleset visualization analysis tool based on segmentation. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062196

[45] M. Kolomeec, G. Gonzalez-Granadillo, E. Doynikova, A. Chechulin, I. Kotenko, and H. Debar. Choosing models for security metrics visualization. In *International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security*, pp. 75–87. Springer, 2017.

[46] E. Krokos, A. Rowden, K. Whitley, and A. Varshney. Visual analytics for root dns data. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709205

[47] J. T. Langton and A. Baker. Information visualization metrics and methods for cyber security evaluation. *2013 IEEE International Conference on Intelligence and Security Informatics*, pp. 292–294, 2013.

[48] B. Laughlin, C. Collins, K. Sankaranarayanan, and K. El-Khatib. A visual analytics framework for adversarial text generation. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161563

[49] L. Leichtnam, E. Totel, N. Prigent, and L. Me. Starlord: Linked security data exploration in a 3d graph. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2017. doi: 10.1109/VIZSEC.2017.8062203

[50] S. Liu and W. He. Research on security visualization: A survey. In *The 59th Annual IACIS Conference*, 2019.

[51] A.-P. Lohfink, S. D. D. Anton, H. D. Schotten, H. Leitte, and C. Garth. Security in process: Visually supported triage analysis in industrial process data. *IEEE Transactions on Visualization and Computer Graphics*, 26(4):1638–1649, 2020. doi: 10.1109/TVCG.2020.2969007

[52] R. Marty. *Applied security visualization*. Addison-Wesley Professional, 2008.

[53] J. McClain, A. Silva, G. Emmanuel, B. Anderson, K. Nauer, R. Abbott, and C. Forsythe. Human performance factors in cyber security forensic analysis. *Procedia Manufacturing*, 3:5301–5307, 2015.

[54] J. Muchagata and A. Ferreira. How can visualization affect security? In *ICEIS (2)*, pp. 503–510, 2018.

[55] S. Mukkamala and A. H. Sung. Identifying significant features for network forensic analysis using artificial intelligent techniques. *International Journal of digital evidence*, 1(4):1–17, 2003.

[56] T. Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009. doi: 10.1109/TVCG.2009.111

[57] A. Nadeem, S. Verwer, and S. J. Yang. Sage: Intrusion alert-driven attack graph extractor. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 36–41, 2021. doi: 10.1109/VizSec53666.2021.00009

[58] A. P. Norton and Y. Qi. Adversarial-playground: A visualization suite showing how adversarial examples fool deep learning. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2017. doi: 10.1109/VIZSEC.2017.8062202

[59] R. Ošlejšek, V. Rusňák, K. Burská, V. Švábenský, and J. Vykopal. Visual feedback for players of multi-level capture the flag games: Field usability study. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–11, 2019. doi: 10.1109/VizSec48167.2019.9161386

[60] S. O'Shaughnessy. Image-based malware classification: A space filling curve approach. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161583

[61] C. Peng, D. Schwartz, D. Johnson, B. Stackpole, C. Weeden, J. Marcovecchio, D. Richards, C. Fogle, C. Brown, and V. Walrond. Visualization for spectators in cybersecurity competitions. In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 21–24, 2020. doi: 10.

1109/VizSec51108.2020.00009

[62] S. Peryt, J. Andre Morales, W. Casey, A. Volkmann, B. Mishra, and Y. Cai. Visualizing a malware distribution network. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2016. doi: 10.1109/VIZSEC.2016.7739585

[63] T. Post, T. Wischgoll, A. R. Bryant, B. Hamann, P. Müller, and H. Hagen. Visually guided flow tracking in software-defined networking. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2016. doi: 10.1109/VIZSEC.2016.7739586

[64] S. L. Reynolds, T. Mertz, S. Arzt, and J. Kohlhammer. User-centered design of visualizations for software vulnerability reports. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 68–78, 2021. doi: 10.1109/VizSec53666.2021.00013

[65] J. C. Roberts. State of the art: Coordinated amp; multiple views in exploratory visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, pp. 61–71, 2007. doi: 10.1109/CMV.2007.20

[66] R. Romero-Gomez, Y. Nadji, and M. Antonakakis. Towards designing effective visualizations for dns-based network threat analysis. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062201

[67] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604–1613, 2014. doi: 10.1109/TVCG.2014.2346481

[68] G. R. Santhanam, B. Holland, S. Kothari, and J. Mathews. Interactive visualization toolbox to detect sophisticated android malware. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062197

[69] A. Schreiber, T. Sonnekalb, and L. v. Kurnatowski. Towards visual analytics dashboards for provenance-driven static application security testing. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 42–46, 2021. doi: 10.1109/VizSec53666.2021.00010

[70] M. Schufrin, S. L. Reynolds, A. Kuijper, and J. Kohlhammer. A visualization interface to improve the transparency of collected personal data on the internet. In *2020 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2020. doi: 10.1109/VizSec51108.2020.00007

[71] A. Sethi and G. Wills. Expert-interviews led analysis of eevi — a model for effective visualization in cyber-security. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062195

[72] H. Shiravi, A. Shiravi, and A. A. Ghorbani. A survey of visualization systems for network security. *IEEE Transactions on Visualization and Computer Graphics*, 18(8):1313–1329, 2012. doi: 10.1109/TVCG.2011.144

[73] H. Siadati, B. Saket, and N. Memon. Detecting malicious logins in enterprise networks using visualization. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739582

[74] A. Sopan and K. Berlin. Ai total: Analyzing security ml models with imperfect data in production. In *2021 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 10–14, 2021. doi: 10.1109/VizSec53666.2021.00006

[75] A. Sopan, M. Berninger, M. Mulakaluri, and R. Katakam. Building a machine learning model for the soc, by the input from the soc, and analyzing it for the soc. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709231

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

[77] S. S. Subramanian, P. Pushparaj, Z. Liu, and A. Lu. Explainable visualization of collaborative vandal behaviors in wikipedia. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–5, 2019. doi: 10.1109/VizSec48167.2019.9161504

[78] M. Syamkumar, R. Durairajan, and P. Barford. Bigfoot: A geo-based visualization methodology for detecting bgp threats. In *2016 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2016. doi: 10.1109/VIZSEC.2016.7739583

[79] R. Tamassia, B. Palazzi, and C. Papamanthou. Graph drawing for security visualization. In I. G. Tollis and M. Patrignani, eds., *Graph Drawing*, pp. 2–13. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[80] The VizSec team. The IEEE Symposium on Visualization for Cyber Security (VizSec). `https://vizsec.org/`, 2022. Accessed: 20.6.2022.

[81] R. Theron, R. Magán-Carrión, J. Camacho, and G. M. Fernndez. Network-wide intrusion detection supported by multivariate analysis and interactive visualization. In *2017 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2017. doi: 10.1109/VIZSEC.2017.8062198

[82] K. Tuma, G. Calikli, and R. Scandariato. Threat analysis of software systems: A systematic literature review. *Journal of Systems and Software*, 144:275–294, 2018.

[83] A. Ulmer, M. Schufrin, D. Sessler, and J. Kohlhammer. Visual-interactive identification of anomalous ip-block behavior using geo-ip data. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–8, 2018. doi: 10.1109/VIZSEC.2018.8709182

[84] A. Ulmer, D. Sessler, and J. Kohlhammer. Netcapvis: Web-based progressive visual analytics for network packet captures. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–10, 2019. doi: 10.1109/VizSec48167.2019.9161633

[85] M. Varga, C. Winkelholz, and S. Träber-Burdin. An exploration of cyber symbology. In *2019 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–5, 2019. doi: 10.1109/VizSec48167.2019.9161577

[86] M. Wagner, F. Fischer, R. Luh, A. Haberson, A. Rind, D. A. Keim, and W. Aigner. A survey of visualization systems for malware analysis. In *EuroVis*, 2015.

[87] Y. Yang, J. Collomosse, A. K. Manohar, J. Briggs, and J. Steane. Tapestry: Visualizing interwoven identities for trust provenance. In *2018 IEEE Symposium on Visualization for Cyber Security (VizSec)*, pp. 1–4, 2018. doi: 10.1109/VIZSEC.2018.8709236

[88] T. Zhang, X. Wang, Z. Li, F. Guo, Y. Ma, and W. Chen. A survey of network anomaly visualization. *Science China Information Sciences*, 60(12):121101, 2017.

[89] Y. Zhang, Y. Xiao, M. Chen, J. Zhang, and H. Deng. A survey of security visualization for computer network logs. *Security and Communication Networks*, 5(4):404–421, 2012.

[90] C. Zhong, T. Lin, P. Liu, J. Yen, and K. Chen. A cyber security data triage operation retrieval system. *Computers & Security*, 76:12–31, 2018.