Heliyon 9 (2023) e13025

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Enhancing cybersecurity situation awareness through visualization: A USB data exfiltration case study

Mu-Huan (Miles) Chung [a,*], Yuhong (Alisha) Yang [b], Lu Wang [a], Greg Cento [b], Khilan Jerath [b], Parwinder Taank [b], Abhay Raman [b], Jonathan H. Chan [c], Mark H. Chignell [a]

[a] *Mechanical and Industrial Engineering, University of Toronto, 5 King's College Rd, Toronto, M5S 3G8, ON, Canada*
[b] *Sun Life Financial Inc, 1 York St., Toronto, M5J 0B6, ON, Canada*
[c] *Innovative Cognitive Computing (IC2) Research Center, King Mongkut's University of Technology Thonburi, 126 Pracha Uthit Rd, Bang Mot, Thung Khru, Bangkok, 10140, Thailand*

ARTICLE INFO

Keywords:
Security visualization
Situation awareness
Data exfiltration
Collaborative dashboard

ABSTRACT

Employees who have legitimate access to an organization's data may occasionally put sensitive corporate data at risk, either carelessly or maliciously. Ideally, potential breaches should be detected as soon as they occur, but in practice there may be delays, because human analysts are not able to recognize data exfiltration behaviors quickly enough with the tools available to them. Visualization may improve cybersecurity situation awareness. In this paper, we present a dashboard application for investigating file activity, as a way to improve situation awareness. We developed this dashboard for a wide range of stakeholders within a large financial services company. Cybersecurity experts/analysts, data owners, team leaders/managers, high level administrators, and other investigators all provided input to its design. The use of a co-design approach helped to create trust between users and the new visualization tools, which were built to be compatible with existing work processes. We discuss the user-centered design process that informed the development of the dashboard, and the functionality of its three inter-operable monitoring dashboards. In this case three dashboards were developed covering high-level overview, file volume/type comparison, and individual activity, but the appropriate number and type of dashboards to use will likely vary according to the nature of the detection task). We also present two use cases with usability results and preliminary usage data. The results presented examined the amount of use that the dashboards received as well as measures obtained using the Technology Acceptance Model (TAM). We also report user comments about the dashboards and how to improve them.

1. Introduction

While there have been considerable advances in hardening perimeters and protecting sensitive data from external attacks, it is more challenging for organizations to protect against data loss from the actions of insiders (who may be careless or malicious but damaging in either case). In one survey, around 27% of cybercrime was conducted by insiders (either employees or people masquerading as

* Corresponding author.
E-mail address: mhmchung@gmail.com (M.-H.(M. Chung).

https://doi.org/10.1016/j.heliyon.2023.e13025
Received 23 September 2022; Received in revised form 12 January 2023; Accepted 13 January 2023
Available online 16 January 2023

employees), with the results of insider attacks being potentially more severe than outside attacks [1].

Employees have a legitimate need to use data as part of their work role and it is often challenging to identify illegitimate, but infrequent, data usage. Due to the potential harm of data exfiltration, administrators should quickly detect and investigate potentially harmful employee activity. Detecting problems with insiders is particularly challenging since harmful activities may be performed by employees with legitimate access. Current insider anomaly detection tools typically [2,3] use one of three approaches 1) automated machine learning (ML); 2) data visualization and interactive dashboards; 3) a combination of approaches 1 and 2.

Given the threats and challenges, organizations need tools that can improve their situation awareness of how employees are transferring confidential and proprietary files, as part of a larger threat detection process. The most common ML approach for threat detection uses statistics and mathematical models to detect any abnormal deviation from the normal activity records based on single-point estimates of threat likelihood [4]. Activity-based detection requires a clear detection goal or a prediction target variable to train and evaluate the result [5]. However, employee behavior frequently changes due to factors such as daily workload, individual working style, access availability, and the nature of the task being performed. Activity-based models can only flag events as potentially malicious, resulting in high levels of false positive errors [6,7]. While the use of rules and heuristics may reduce the number of false positives, it is difficult to keep rule sets up to date in response to a rapidly changing threat environment [6–8].

ML algorithms work best with well-ordered, continuous, and non-sparse data. However, log files relevant to data exfiltration typically contain a lot of textual information (e.g., Device and Path Names), but include relatively few continuous variables other than file size), which can constrain the performance of an ML algorithm. For data transfers to external storage devices (e.g., copying to USB devices), limited data in log files may make automated ML system development infeasible or unreliable.

Data visualization methods are used by many enterprises for daily activity monitoring and detection. Visualizations can provide overviews of large amounts of data in the cybersecurity domain [9]. While visualization may not be needed for events that are easy to label automatically as suspicious (e.g., an unauthorized user attempted to log in to a sensitive database), it is potentially useful for monitoring USB file copying. This is because situations that look anomalous may arise for many reasons, most being legitimate. Thus, with respect to USB copying, visualization can be used for exploration and verification, and for the collection of supporting evidence [10].

The main difficulty with existing security visualization strategies is poor collaborative decision-making [11]. For instance, current strategies do not consider variations in data usage by people in different roles, or with different missions, and they tend to be designed for skilled analysts, rather than Field Domain Specialists (FDSs) such as product managers. The lack of collaboration among those with potentially relevant knowledge increases the possibility of missing potentially harmful patterns of behavior that signal anomalies worthy of investigation.

In designing an insider file copying activity monitoring dashboard, our overall goal was to provide assistive tools to generate insights that facilitate informed decision-making. The methods that we developed for visualizations to monitor file copying should also be beneficial in the development of other security tools.

2. Related work

In this section, we discuss useable security visualization, noting the deficiencies of current security dashboard design. We review past research on the importance of situation awareness in cybersecurity and previous file transfer visualization approaches supporting anomalous behavior detection.

2.1. Situation awareness in cybersecurity

Situational awareness is traditionally defined as “the perception of the surroundings and derivative implications critical to decision-makers in complex, dynamic areas such as military command and security” [12]. Maximizing situational awareness may guarantee “operational risks to be mitigated, managed, or resolved prior to a mission or during operations” [13]. Barford et al. [14] introduced the term “cyber situational awareness”, to refer to the application of situational awareness in cybersecurity. There are seven major requirements that describe what stakeholders should be aware of to increase their awareness and to make their cyber network safe (we consider the following four requirements relevant to our concern with the construction of detection models).

- Awareness of the **current situation** (also known as situation perception): This includes identification of where the attack is coming from, who is conducting the attack, what vectors are used, which vulnerabilities are being exploited, and what the targets are.
- Awareness of **adversary behavior** (the trend of the attack): This may help understand the motivation and targets of the adversary. This information can be gathered from logs and endpoint behavioral data.
- Awareness of **quality and trustworthiness** (of the collected situation awareness information items and the knowledge-intelligence decisions derived from these information items).
- Awareness of plausible **future evolution** (of the current situation): This involves the prediction function of ML models.

In this paper we examine the role/usability of visualization dashboards in addressing the current problem of low situation awareness in cybersecurity applications.

2.2. *Useable security visualization*

Accessible and useable security system design should be informative, reliable, interpretable, assistive, functional, navigable, responsive, adaptive, and easy to learn [15]. Whitten [16] claimed that a useable security tool should allow users to 1) be aware of tasks they need to perform, 2) perform the task 3) avoid dangerous errors and 4) be comfortable with using it again. A useable security product should also enable users to make informed decisions in a specific business context [17,18]. Staheli [19] discussed some dimensions for evaluating security visualization for human-machine collaborative systems, including user experience and preference, usability and learnability, effect on collaboration, insights generation, task demands, cognitive workload, and component interoperability. In the research reported in this paper we focused on whether the visualization supported a variety of user tasks, and the extent to which the tool had a positive impact on situation awareness in representative contexts.

Security experts and IT consultants typically determine security practice according to the business impacts of security controls, without considering usability, or impacts on users, fully [20]. Design for usability in security visualization has often focused on distancing users from details to avoid the system becoming confusing and overwhelming [17]. However, detailed, and consumable, insights and actionable evidence are needed to aid effective anomaly detection.

2.3. *Visualization for insider file transfer*

Pfleeger et al. [21] defined insiders as people who can legitimately access an organization's network and digital device, and defined insider threats as actions, conducted by an insider, that might put organizations' data or resources at risk.

File, or data, exfiltration involves not just copying or transferring proprietary information by employees, but crucially, the transfer of that information outside the organization. The existence of exfiltration depends on the context since some transfers of information outside the organization may be legitimate. Exfiltration events are particularly salient in the financial industry, where information leakage of personal sensitive data will sabotage consumer trust and violate privacy obligations. Effective internal risk management systems are key to securing safe daily operations within a company. Risk management involves the ability to define malicious cases, detect malicious events, and prevent anomalies from happening [22,23].

Many different types of visualization have been used for cybersecurity data, including line graphs, pie charts, cumulative distribution charts, ranked lists, tables of descriptive statistics, activity timelines, and risk scoring [24]. Pattern-based visualization identifies abnormal activity from comparison with historical records in an interpretable way, and threats may be assessed by examining the latest data with the employee's normal behavior or with the behavior of a relevant comparison group [25,26]. Gamachchi et al. [27] used histograms and frequency line graphs to visualize individual logon-logoff behavior, while Legg et al. [28] used circular heatmaps with color indicators to visualize login frequencies. Hanniel et al. [29] used dot maps to visualize anomalous user activity based on time.

In the following discussion, the design of visualization dashboards is presented, where the focus is on the detection of data exfiltration. For more details on the design of the dashboards see Ref. [30].

3. Requirement analysis

Requirements analysis was carried out using interviews and a focus group as described below. The analysis procedure was approved by the University of Toronto ethics review board (Ethics protocol number 39752).

3.1. *User groups and detection process*

Semi-structured interviews were carried out with four people who have been involved with the manual detection from the security

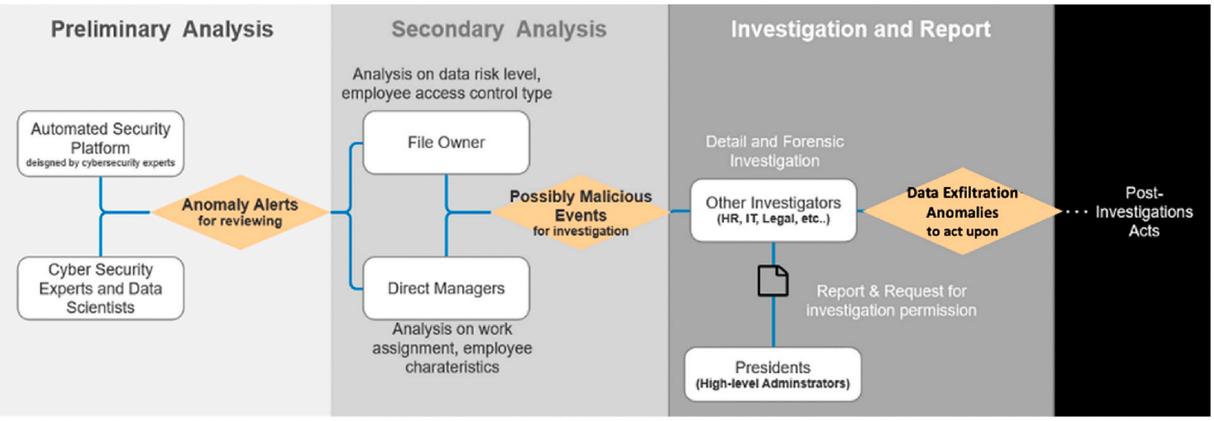

Preliminary Analysis
Secondary Analysis
Investigation and Report
Analysis on data risk level, employee access control type
Automated Security Platform
deisgned by cybersecurity experts
Anomaly Alerts
for reviewing
Cyber Security Experts and Data Scientists
File Owner
Possibly Malicious Events
for investigation
Direct Managers
Analysis on work assignment, employee charateristics
Detail and Forensic Investigation
Other Investigators
(HR, IT, Legal, etc..)
Data Exfiltration Anomalies
to act upon
Post-Investigations Acts
Report & Request for investigation permission
Presidents
(High-level Adminstrators)

Fig. 1. Standard detection process for USB file copying anomalies at the financial company studied in this research (Fig. 2 in Ref. [30]).

analytics team in the company studied to understand the current file activity detection process and the stakeholders. Further details about the interview results are provided by Ref. [30]. They found that there were five main types of stakeholders with respect to USB copying (cybersecurity analysts/experts; file owners; team leaders; high-level administrators; other investigators).

Prior to our research, a decision as to whether an activity currently under scrutiny is a potentially malicious event worthy of more detailed investigations was typically made using the standard processing shown in Fig. 1.

The figure shows a three-stage process where different stakeholders are involved in each of the sequential stages in the process.

3.2. Harmful anomalies and the detection objective

After the initial round of exploratory user research reported in the previous section, the second round of interviews and focus group studies were performed to define detection goals, so as to identify malicious use cases of concern, and design requirements for USB file copying monitoring and detection of abnormal events.

A total of 17 participants participated in the interviews. The questions were co-designed with a cybersecurity expert and verified by the AVP of the security analytics department, and a list of use cases was updated after each interview. The sample included two department directors, two risk division VPs, three team managers, three security data analysts, two non-technical supporting investigators (HR division) and the five people in a team that was involved in developing internal security monitoring and alerts for the financial company.

Interviews identified the following characteristics of file transfers that were regarded as anomalous, e.g., transfers with high spikes in volume and frequency of.

- A series of semantically related sensitive files
- Timestamps outside of working hours (less effective as an indicator when employees working from home)
- Locations involving more than one computer
- Irregular behavior in comparison with other colleagues at the same job position and seniority (Time of events, volume of transfer, type of file transfer)
- Irregular behavior in comparison with other colleagues in the same team or branch (Time of events, volume of transfer, type of file transfer)
- Irregular behavior in comparison with the person's own historical records (Time of events, volume of transfer, USB device, type of file transfer)

Based on their tacit knowledge about human behaviour and job roles, the interviewees also recommended that the following groups of employees with out-of-date USB transfer access should receive particular attention.

- Employees near their employment termination date
- Employees with administrative control access
- Employees with special employment type (agent, temporary, casual, vendor, contractor, fixed term)

4. Dashboard design

4.1. Interactive filters

The process of designing the visualization dashboards is described by Ref. [30]. This process included the design of interactive filters that made it easy to define customized views of the data. These filters can be identified as follows.

1. Employee Name/ID: Employee Identification
2. Business Group: Department and associate region
3. Job Role: Work position title and seniority
4. Manager: Team lead by a specific leader
5. Employee Type: Type of employment (e.g., regular, temporary, casual, fixed term, etc.)
6. Employee Status: Whether the employee is active or inactive
7. Admin: Whether the employee has full administrative access to their company laptop
8. GCE: Whether the employee has access to government-related data
9. USB Exemption: Type of USB access (e.g., read write, transfer, etc.)
10. Termination: whether the employee will leave the company within 90 days
11. High-Risk File Type: True or False filter, where high-risk categories are defined by the company including all Microsoft file types, data files, zip files, etc.
12. Outside Working Hours: Whether the activity is done outside the working hours/date

The 12 filters that were developed helped different users carry out the tasks relevant to them. For example, using the manager filter, the team leader could view all the employees within a team and detect abnormal activities by comparing the characteristics of a possibly anomalous event against the background activities of other team members. The assumption and rationale behind each filter in the dashboard can be found in Appendix A of this paper.

Additionally, analysts could combine filters to filter out activities with a higher risk, before alerting other stakeholders. An example of high-level (aggregated) filtering of this sort is using filters #5, #11, and #12 in combination to filter out a) "temporary" employees; b) those who extracted a high volume of files outside working hours; c) one week before the termination date.

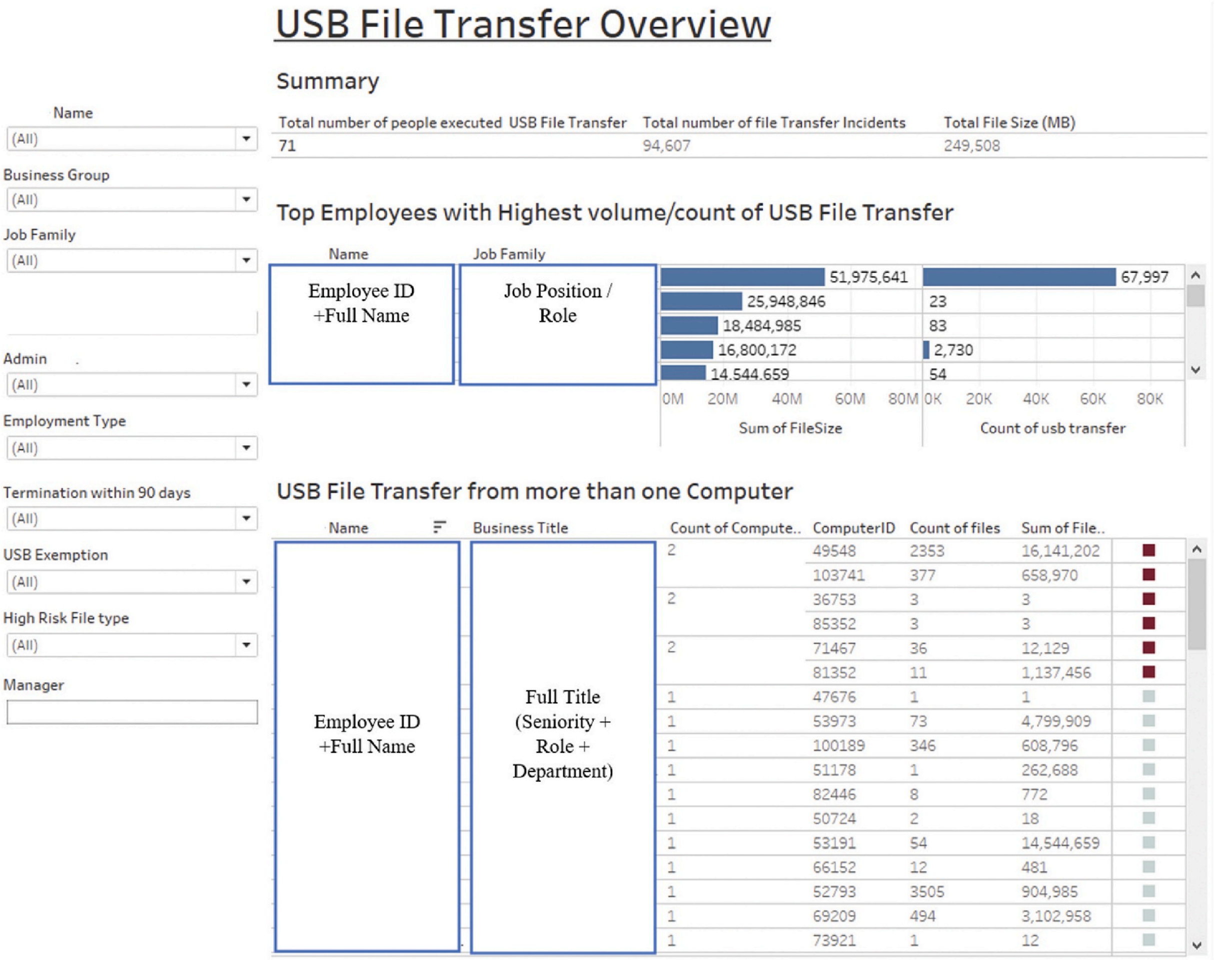
4.2. Anomaly investigation dashboards

The tool we developed includes 3 interoperable dashboards that highlight suspicious characteristics that may be used for supporting reliable anomaly detection.

The first dashboard, illustrated in Fig. 2, shows a high-level overview of USB file copying events customized by the filter inputs over a selected period. The core value of the dashboard is that it illustrates an aggregated and customized view of file transfer status over a defined period, allowing analysts to identify suspicious activities that might need further investigation using the other two dashboards.

The first view at the top of Fig. 2 is a "Summary of USB file transfer activities", which shows the total number of employees who have copied files to USB, along with the total count of files, and the total file size. The second view is an ordered bar chart that illustrates the aggregated file transfer file size and frequencies for an employee. The role of the employee is also shown, since it may be informative. For example, it might be reasonable for a conference-attending employee to extract a lot of files for event preparation, but it may be suspicious if a junior financial analyst ranked high in the chart (without a similarly obvious reason). The third view uses a red dot indicator to identify employees who transferred files from more than one computer with their corresponding business tile, computer IDs, count of files and sum of file size extracted from each computer.

As shown in Fig. 3, the second dashboard allowed users to examine the volumes of USB file transfers. More details about this dashboard can be found in Ref. [30]. One of the features of this dashboard that is noted here is that color coding is used to indicate the degree of anomaly (e.g., a red circle indicates a value that is 3 or more standard deviations from the mean of the property being viewed).



USB File Transfer Overview

Name: (All)

Business Group: (All)

Job Family: (All)

Admin: (All)

Employment Type: (All)

Termination within 90 days: (All)

USB Exemption: (All)

High Risk File type: (All)

Manager:

Summary

Total number of people executed USB File Transfer	Total number of file Transfer Incidents	Total File Size (MB)
71	94,607	249,508

Top Employees with Highest volume/count of USB File Transfer

Name	Job Family	Sum of FileSize	Count of usb transfer
Employee ID +Full Name	Job Position / Role	51,975,641	67,997
		25,948,846	23
		18,484,985	83
		16,800,172	2,730
		14,544,659	54

USB File Transfer from more than one Computer

Name	Business Title	Count of Compute..	ComputerID	Count of files	Sum of File..	
Employee ID +Full Name	Full Title (Seniority + Role + Department)	2	49548	2353	16,141,202	🟥
			103741	377	658,970	🟥
		2	36753	3	3	🟥
			85352	3	3	🟥
		2	71467	36	12,129	🟥
			81352	11	1,137,456	🟥
		1	47676	1	1	⬜
		1	53973	73	4,799,909	⬜
		1	100189	346	608,796	⬜
		1	51178	1	262,688	⬜
		1	82446	8	772	⬜
		1	50724	2	18	⬜
		1	53191	54	14,544,659	⬜
		1	66152	12	481	⬜
		1	52793	3505	904,985	⬜
		1	69209	494	3,102,958	⬜
		1	73921	1	12	⬜

Fig. 2. The USB file transfer overview dashboard.

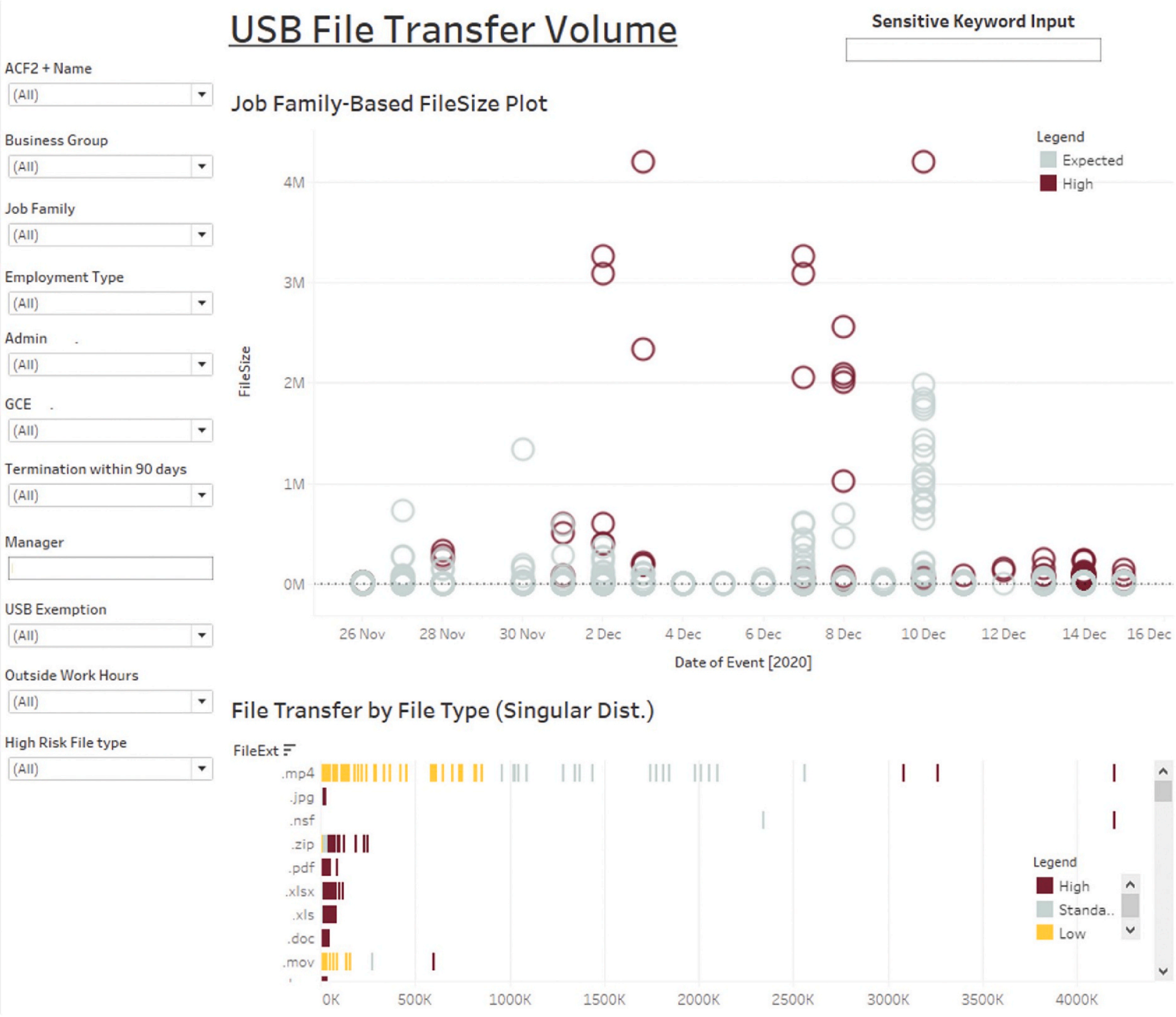
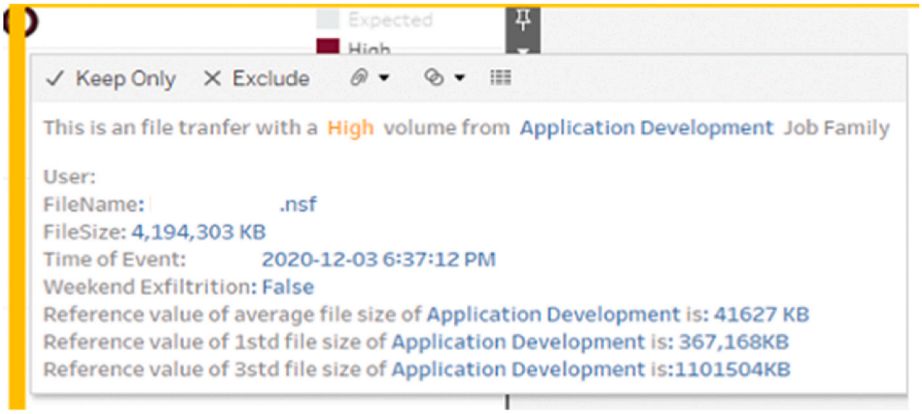


Fig. 3. The USB file transfer volume dashboard (Fig. 3 in Ref. [30]).



This is an file tranfer with a High volume from Application Development Job Family

User:

FileName: .nsf

FileSize: 4,194,303 KB

Time of Event: 2020-12-03 6:37:12 PM

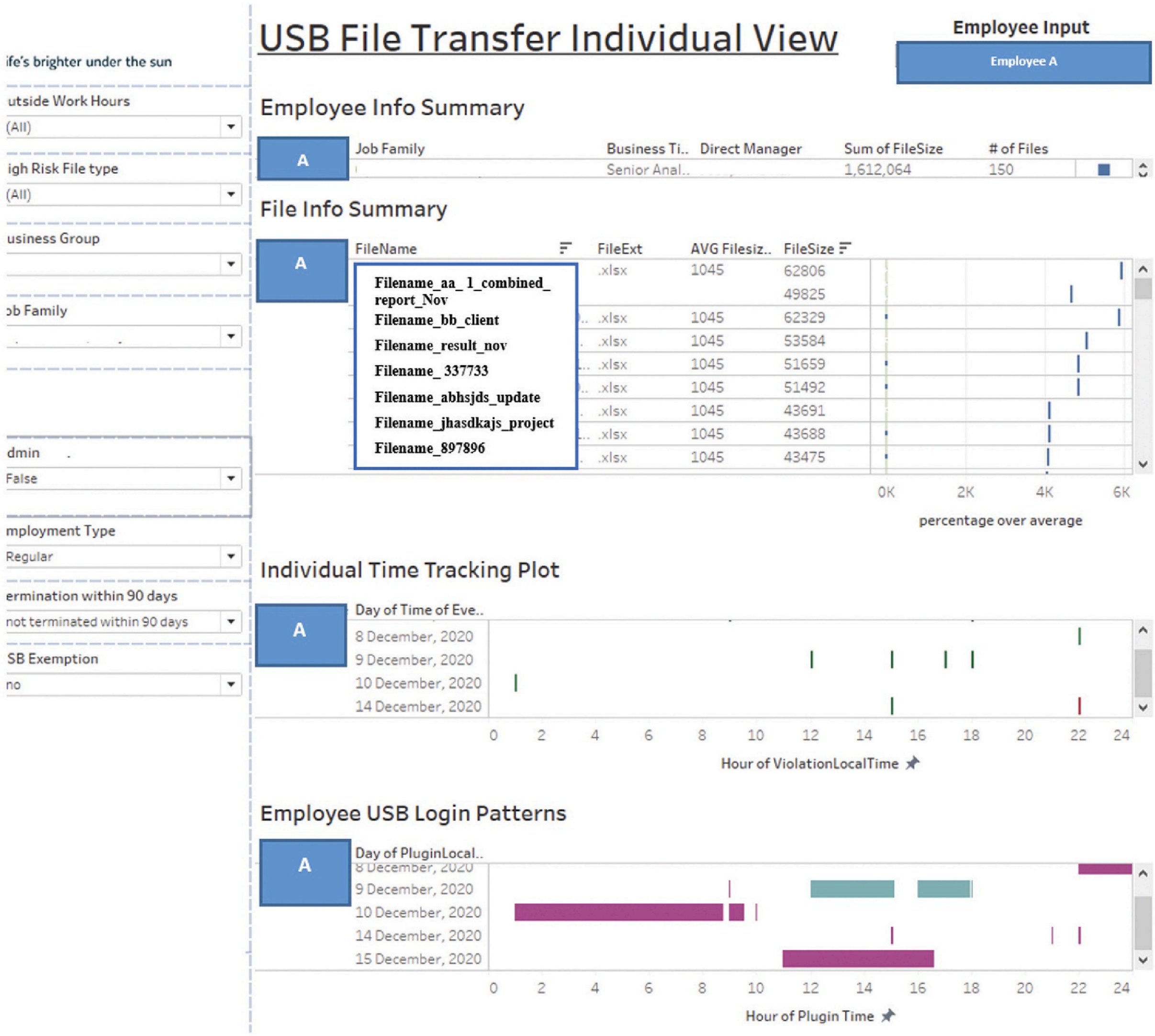
Weekend Exfiltrition: False

Reference value of average file size of Application Development is: 41627 KB

Reference value of 1std file size of Application Development is: 367,168KB

Reference value of 3std file size of Application Development is:1101504KB

Fig. 4. Tooltip of the Job-family based File-size plot.



ife's brighter under the sun

utside Work Hours: (All)

igh Risk File type: (All)

usiness Group:

ob Family:

dmin: False

mployment Type: Regular

ermination within 90 days: not terminated within 90 days

SB Exemption: no

USB File Transfer Individual View

Employee Input: Employee A

Employee Info Summary

Job Family	Business Ti..	Direct Manager	Sum of FileSize	# of Files
	Senior Anal..		1,612,064	150

File Info Summary

Filename_aa_ 1_combined_ report_Nov
Filename_bb_client
Filename_result_nov
Filename_ 337733
Filename_abhsjds_update
Filename_jhasdkajs_project
Filename_897896

FileExt	AVG Filesiz..	FileSize
.xlsx	1045	62806
		49825
.xlsx	1045	62329
.xlsx	1045	53584
.xlsx	1045	51659
.xlsx	1045	51492
.xlsx	1045	43691
.xlsx	1045	43688
.xlsx	1045	43475

Individual Time Tracking Plot

Employee USB Login Patterns

Fig. 5. Dashboard of USB file transfer individual view.

The final two dashboards shown here (Figs. 4 and 5) focus on job roles (Fig. 4) and individuals (Fig. 5), respectively. When the user hovers over a circle or bar in Fig. 4 a tooltip appears. The tooltip provides specific information related to this incident, including username, job role, file name, and file size. For comparison purposes the final three lines in the display show statistical reference values. Tooltip information boxes are embedded in all of the dashboards, when applicable, providing additional information intended to help users make quicker and more reliable decisions.

The dashboard that focuses on individuals (Fig. 5) includes 4 plots and a user input function. The Employee Information Summary includes username, job role, business tile, direct managers, and the basic statistics of USB file transfers. The File Info Summary plot displays all the files transferred by an individual over a selected period and shows a percentage over the average value to visualize how far incidents deviate from the average file size for each file type. An individual time tracking plot shows file transfers against the time of day, with color coding indicating how atypical the size of file transfer is for a particular time of day. The Employee USB Login Pattern graph plots plugin duration with different colors indicating different USB devices used.

5. Dashboard use cases

The following examples show how the designed dashboards can be used to provide interpretable information that enables investigators to make reliable decisions in a short time. Interaction and visualization design are easy to navigate and interpret, which allows all defined stakeholders to access the information without needing data analytic skills.

5.1. Anomaly detection

Our first evaluation used interviews and anecdotal observations to provide qualitative feedback on how the dashboards were perceived. One of the key issues was how easy the dashboards would be to understand and use for non-technical staff. The general feedback was that the dashboards were easy to understand and use. It was felt that the visualization reports were relatively intuitive to understand and could potentially reduce the amount of time necessary to investigate potential anomalies. Based on this promising initial feedback a quantitative evaluation of dashboard use and acceptance was carried out, as reported in Section 6 of this paper.

In this subsection, an example is provided of how the dashboard can support true anomaly detection, in a short period of time, with the collaboration of security analysts and team managers.

Security analysts could use the USB File Transfer Overview Dashboard to monitor high-level patterns and to identify employees with high spikes of file copying. For instance, Fig. 6 shows how employee “Ann”, from the quality assurance team, ranked first in the volume count plot across the company. The value is significantly higher than any other employees ranked, and there are no other employees from the same job family ranked highly in this chart. Therefore, “Ann” is worth reviewing using another dashboard, as shown in Fig. 7.

Managers can use the File transfer Volume dashboard (shown in Fig. 7) to visualize the daily activities of “Ann”. The figure shows the view that results after selecting the employee using a filter combined with a search on the term “result” (entered into the search field shown in the top right of the figure). As shown in Fig. 7, high volumes of copying occurred during weekends, and all the files are “result” files from a special project. Based on the information presented in this visualization, and the implicit knowledge of managers, it appears that the copying activities of “Ann” are anomalous and potentially problematic.

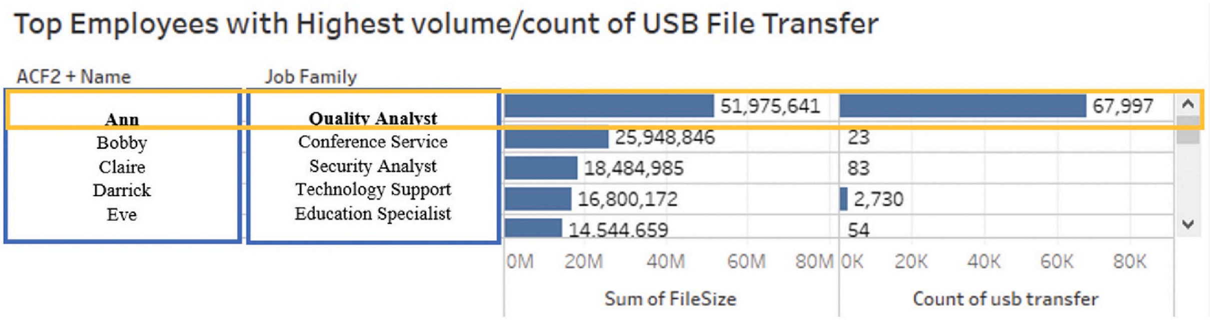
5.2. False positive screening

Anomaly detection tends to have a high rate of false positives because outliers are detected based on one variable at a time, and thus many outliers are detected when there are a high number of variables. As shown in the following example, dashboards should help to reduce the number of false positives by including more potential users in the validation process and by using contextual information visualization to reduce the time and effort needed to screen false positives.

When security analysts want to monitor the file activity of irregular employees, (e.g., Temporary, Contractors, Vendor, Fixed Termed, etc.), they can use the employment role/family filters in each dashboard to customize the view. In the example shown in Figs. 8 and 9, an analyst discovered a temporary employee, “Bob” who had a high frequency of USB activities and who copied many files with similar filenames. If only technical analysts were investigating, they might assume that the activity was malicious. However, if the direct manager reviewed this information, she could check if the risk level of file transfers was consistent with the work being performed by Bob. In this case, employee “Bob” was about to go on an international business trip (a fact known to his managers), and he requested temporary USB transfer access so that copying files to his PC before attending a conference was seen to be appropriate, and not a problematic anomaly.

6. Evaluation

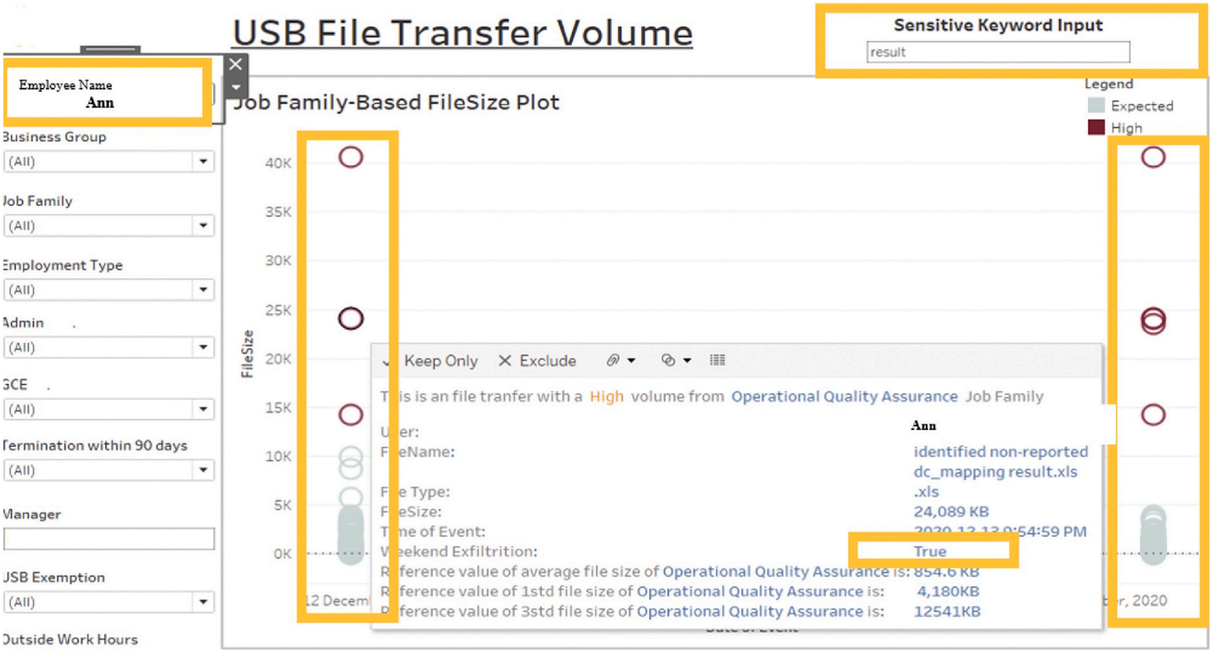
After developing the dashboards through iterative dashboard design and testing within the design team, we carried out a task-based evaluation, followed by an evaluation of the dashboards with other stakeholders. The task-based evaluation of testing protocols is shown in Appendix B. Since the dashboard was designed for monitoring internal security problems, the number of target users was relatively low, and it was hard to carry out observations with them due to their busy schedules. Therefore, we designed a relatively low-impact 2-week remote pilot testing protocol based on the Technology Acceptance Model (TAM) [31]. The testing procedure included 3 stages, 1 h demo presentation, 2-week remote testing, and a feedback debrief. We followed this with usage observations after the implementation of our tool.



Top Employees with Highest volume/count of USB File Transfer

ACF2 + Name	Job Family	Sum of FileSize	Count of usb transfer
Ann	**Quality Analyst**	51,975,641	67,997
Bobby	Conference Service	25,948,846	23
Claire	Security Analyst	18,484,985	83
Darrick	Technology Support	16,800,172	2,730
Eve	Education Specialist	14,544,659	54

Fig. 6. Number and Files Sizes of USB File Transfers with most active Employee “Ann” highlighted for transfer. The plot again focuses on atypical behavior, relative to job roles and time of day.



USB File Transfer Volume

Sensitive Keyword Input: result

Employee Name: Ann

Business Group: (All)

Job Family: (All)

Employment Type: (All)

Admin: (All)

GCE: (All)

Termination within 90 days: (All)

Manager:

USB Exemption: (All)

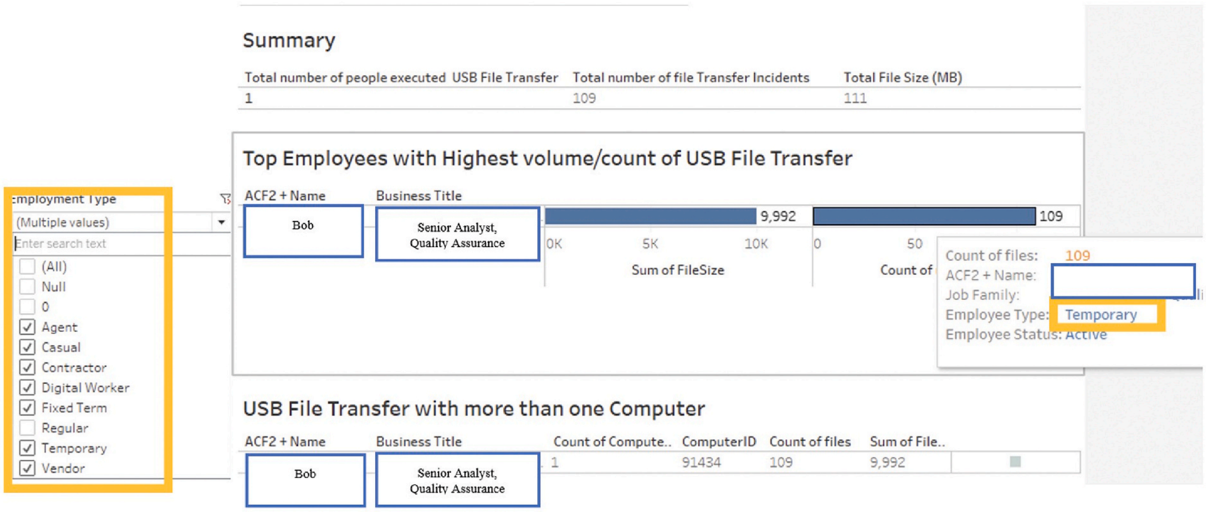
Outside Work Hours

Job Family-Based FileSize Plot

This is an file tranfer with a High volume from Operational Quality Assurance Job Family

User:	Ann
FileName:	identified non-reported dc_mapping result.xls
File Type:	.xls
FileSize:	24,089 KB
Time of Event:	2020-12-13 9:54:59 PM
Weekend Exfiltrition:	True
Reference value of average file size of Operational Quality Assurance is:	854.6 KB
Reference value of 1std file size of Operational Quality Assurance is:	4,180KB
Reference value of 3std file size of Operational Quality Assurance is:	12541KB

Fig. 7. Plot of file transfers on different days showing that employee "Ann" copied to USB a lot of files containing "result" in the file name, on weekends.



Summary

Total number of people executed USB File Transfer	Total number of file Transfer Incidents	Total File Size (MB)
1	109	111

Top Employees with Highest volume/count of USB File Transfer

ACF2 + Name	Business Title	Sum of FileSize	Count of
Bob	Senior Analyst, Quality Assurance	9,992	109

Count of files: 109
ACF2 + Name:
Job Family:
Employee Type: Temporary
Employee Status: Active

Employment Type: (Multiple values)

- ☐ (All)
- ☐ Null
- ☐ 0
- ☑ Agent
- ☑ Casual
- ☑ Contractor
- ☑ Digital Worker
- ☑ Fixed Term
- ☐ Regular
- ☑ Temporary
- ☑ Vendor

USB File Transfer with more than one Computer

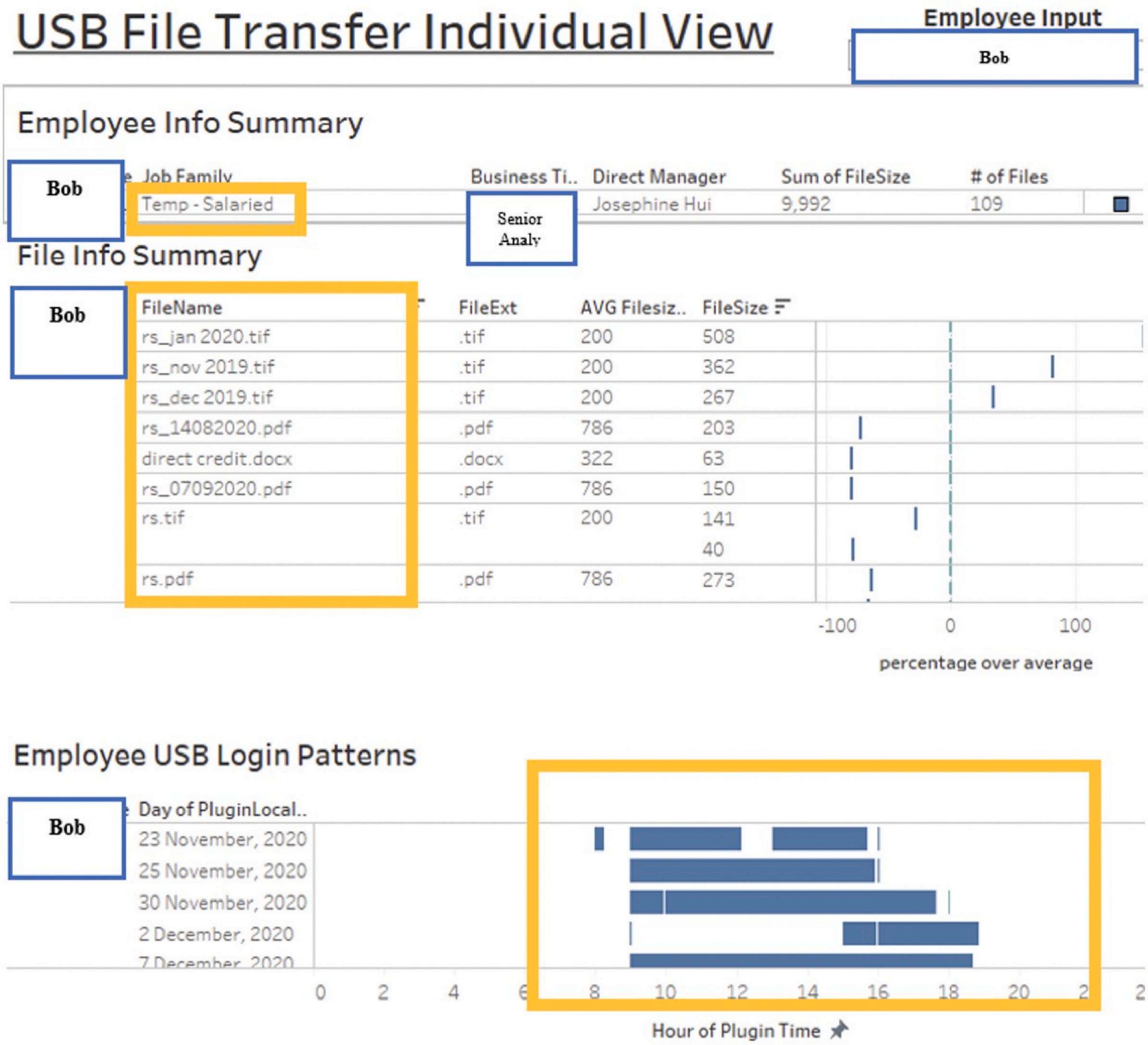
ACF2 + Name	Business Title	Count of Compute..	ComputerID	Count of files	Sum of File..
Bob	Senior Analyst, Quality Assurance	1	91434	109	9,992

Fig. 8. "Bob" is a non-regular (Temporary) employee who copied files to his personal USB Device.

6.1. Evaluation methodology

We evaluated the dashboards' usability, and the intention to use (for target stakeholders) using TAM. We adopted the adjusted TAM form of this popular methodology and added system compatibility into the model in evaluating perceived usefulness (PE), perceived ease of use (PEOU), and intention to use.

The TAM assumes that user behavioral intentions to use a system are determined by two main factors, a) perceived usefulness; b) perceived ease of use [32,33]. Perceived usefulness indicates whether users believe that the system can help them perform their tasks, and ease of use indicates whether users are able to learn and use the technology without too much effort [32]. These two constructs are affected by external variables and by internal motivation. The determination of specific constructs is assumed to depend on the context of the research [34,35]. Aside from predicting whether the user will use the designed services, TAM can also be used to uncover the reasons for technology adoption and the perceived risk of accepting such technology [36–38].



USB File Transfer Individual View

Employee Input: Bob

Employee Info Summary

Bob

Job Family	Business Ti..	Direct Manager	Sum of FileSize	# of Files
Temp - Salaried	Senior Analy	Josephine Hui	9,992	109

File Info Summary

Bob

FileName	FileExt	AVG Filesiz..	FileSize
rs_jan 2020.tif	.tif	200	508
rs_nov 2019.tif	.tif	200	362
rs_dec 2019.tif	.tif	200	267
rs_14082020.pdf	.pdf	786	203
direct credit.docx	.docx	322	63
rs_07092020.pdf	.pdf	786	150
rs.tif	.tif	200	141
			40
rs.pdf	.pdf	786	273

Employee USB Login Patterns

Bob

Day of PluginLocal..
23 November, 2020
25 November, 2020
30 November, 2020
2 December, 2020
7 December, 2020

Fig. 9. Plot of employees having a high rate of USB transfers, showing high spikes of USB copying for Bob.

In the context of internal USB file activity monitoring, perceived usefulness (PE) refers to the degree to which the stakeholders believe the dashboards could improve detection and monitoring performance and efficiency.

We also considered individual factors, social factors and organizational factors that contribute to behavior intention (BE) to adopt the system after deployment. The individual factor that we considered was self-efficacy, i.e., the extent to which the system as designed matched user skillsets and whether they were confident in using the system to find the relevant information or not. The social factor used was subjective norm, which examined the importance of file activity monitoring in their work. Organizational factors included system accessibility and system compatibility with respect to the current workflow. The factors described the sub-dimensions of PE, PEOU, and BE, which attempt to describe why individuals choose to adopt or not adopt our dashboard in performing file activity monitoring and detection tasks.

Specific questions and indicators for each construct were designed according to the definition and design objectives, presented in AppendixC. The questionnaire items used a seven-point Likert scale, from 1 = "Strongly Disagree" to 7 = "Strongly Agree".

6.2. Participants

Participants in the evaluation were selected based on the following inclusion criteria.

1. Never involved in the design process (to avoid designer bias)
2. Authorized to access and analyze employees' personal information and activity records

3. Knowledgeable about the characteristics of the stakeholder group identified in Section 3

Due to the sensitivity of the data used in developing the dashboard, we were only able to include six participants in the evaluation. The participants included one team manager, two security division leaders and three cybersecurity analysts from the "blue" team who were tasked with defending the company against cybersecurity threats. The final questionnaires were submitted anonymously after obtaining consent from both the employee participants, the company and the University Toronto research ethics board.

Group A "Team Manager" represented user groups who had knowledge in work and employees under their supervision, sufficient skills in using digital devices, but limited skills in data management and security analysis • This included user group "Direct Managers", "High-level Administrators", and "Other Investigators"

Group B "Security Division Lead" This included user groups "Direct Managers", "High-level Administrators", and "Other Investigators" which represented user groups who had knowledge in security management, and sufficient skills in digital devices but no knowledge of specific work role or of employee characteristics that were not recorded in company records.

- This included user groups "File Owner" and "High-level Administrators"

Group C "Security Analyst" conducted the security analysis and monitored the internal employees' file activity currently. They were proficient in security analysis, data analytic and security visualization but had little knowledge of the work roles.

- This included the user group "Security Analysts"

6.3. Evaluation procedure

In stage one of the evaluation, the explanatory demo session was held, where we described and demonstrate the scope, motivation, functionality and use case of the dashboards. The demo was followed by an interview that focused on the following two issues.

- Usability of the design and deficiencies existed in the current design
- Feasibility of putting the technology into their current security analysis workflow

After discussing the issues, the interviewees were asked to suggest security problems/domains that can be solved using visualization tools.

In stage two of the evaluation, the dashboard was published on the internal Tableau server. All participants received the Tableau server link of the dashboard with the user guide for a remote 2-week pilot testing. At the end of remote testing, they were asked to submit the TAM questionnaire (shown in Appendix C), which was designed based on the Adjusted Technology Acceptance Model, with some adjustments (as is typically done with the TAM) to customize the questions for the specific technology use under consideration. In stage three of the evaluation, the final round of evaluation collected feedback on user experience with the dashboard in the real along with user feedback on why they made their questionnaire responses.

7. Results

The interview results (stage one evaluation) indicated that participants understood the dashboard visualization and navigation. They believed that they could perform the monitoring task (Appendix B) and identify malicious events using the dashboard. In addition, potential design deficiencies were identified along with solid recommendations for future changes.

Participants suggested that some insights generated from one dashboard should be populated into the next dashboard automatically. For instance, filter settings should be applied uniformly to all dashboards once an entry has been made; and, if a user identified an employee as anomalous in the Overview Dashboard, the information should be reflected/highlighted in the Individual View Dashboard.

Another recommendation was that in-depth ticket analysis should be considered to identify the mismatch between file transfer activities and access request content. For instance, if an employee requests temporary USB transfer access to download files for client presentations, but the actual activity involves extracting ongoing project-related files, alerts should be visually presented to analysts.

Participants also noted that the risk classification of job families might consider other characteristics. Currently, employees with full administrative access and government-related data access are closely monitored using the dashboard, however, people with client personal information access are another risky group, and inappropriate behavior from that group could bring significant damage to consumer trust. The current dashboard only allowed users to identify malicious individuals or activities and group them into risk groups by the Tableau grouping function. A further consideration was that the dashboard should have a function to label the anomaly after analysis so that the labelled data would be available for future automated ML algorithm development.

Participants suggested that some insights generated from one dashboard should be populated into the next dashboard automatically. For instance, filter settings should be applied uniformly to all dashboards once an entry has been made; and, if a user identified an employee as anomalous in the Overview Dashboard, the information should be reflected/highlighted in the Individual View Dashboard.

Another recommendation was that in-depth ticket analysis should be considered to identify the mismatch between file transfer activities and access request content. For instance, if an employee requests temporary USB transfer access to download files for client

Table 1
Defining different aspects of threat modeling.

Construct	Measurement Instrument	Group 1 Mean (S = 1)	Group2 Mean (S = 2)	Group3 Mean (S = 3)
Perceived ease of use	I find the visualization dashboards easy to use (E_1)	**7**	**6.5**	**6.67**
	Learning how to use the visualization dashboards would be easy for me. (E_2)	**5**	**5**	**4.33**
	It is easy to become skillful at using the visualization system. (E_3)	**6**	**5.5**	**6**
Perceived Usefulness	The exploratory visualization dashboard would improve my performance in USB file activity monitoring and anomaly detection. (U_1)	**7**	**6.5**	**6**
	The exploratory visualization dashboard would increase productivity. (U_2)	**3**	**3**	**5.67**
	The exploratory visualization dashboard would enhance my effectiveness at security analysis work. (U_3)	**3**	**3.5**	**5**
	The exploratory visualization dashboard would make it easier to do my work. (U_4)	**4**	**3.5**	**5**
	I found the various functions in the dashboard were well integrated. (U_5)	**7**	**6.5**	**6.33**
Attitude	Anomaly detection and activity monitoring through exploratory visualization system is a good idea. (A_1)	**7**	**7**	**6.33**
	I am positive towards the system and its future implementation. (A_2)	**6**	**5.5**	**6**
Behavioral Intention	I intend to use the exploratory visualization dashboards frequently. (B)	**4**	**3.5**	**4.67**
Self-Efficacy	I feel confident in finding information in the dashboard. (S_1)	**5**	**5.5**	**6**
	I feel confident in identify malicious cases in the dashboards. (S_2)	**6**	**5**	**5**
	I have the necessary skills for using the dashboard without technical support. (S_3)	**6**	**6**	**6.33**
Subjective Norm	Anomaly detection and file activity monitoring is important for me as a security analyst. (N_1)	**N/A**	**4**	**7**
	To improve my security analytical task performance, more effective detection strategy is necessary. (N_2)	**7**	**4.5**	**6**
System Accessibility	I have no difficulty accessing the tableau server to use the dashboards. (SA)	**7**	**7**	**7**
System Compatibility	Using this dashboard fits my current workflow. (C_1)	**3**	**5.5**	**5**
	Using the exploratory visualization system is compatible with most aspects of the detection process for USB file activity monitoring. (C_2)	**3**	**4.5**	**5**
	Using the system fits well with the way that I intend to monitor USB file activity. (C_2)	**6**	**5**	**5**

presentations, but the actual activity involves extracting ongoing project-related files, alerts should be visually presented to analysts.

Participants also noted that the risk classification of job families might consider other characteristics. Currently, employees with full administrative access and government-related data access are closely monitored using the dashboard, however, people with client personal information access are another risky group and inappropriate behavior from that group could bring significant damage to consumer trust and benefits. The current dashboard only allowed users to identify malicious individual or activities and group them into risk group by the Tableau grouping function. A further consideration was that the dashboard should have a function to label the anomaly after analysis so that the labelled data would be available for future automated ML algorithm development.

The results from the seven-point Likert scale questionnaire are shown in Table 1. The mean of each participant group is calculated separately.

As can be inferred from the results shown in Table 1, the dashboard is accessible, understandable, and easy to use for all user groups, and non-tech participants can navigate the dashboards without the presence of assistance (E_1 = [7,6.5,6.67]; U_1 = [7,6.5,6], U_5 = [7,6.5,6.33], SA = [7,7,7]). Interviewees also stated that the tool's functions are well integrated and that visualizations are easy to interpret. Participants held a positive attitude (A_1 = [7,7,6.33], A_2 = [6,5.5,6]) towards using visualization as an effective aiding tool in anomaly detection, and they found it easy to carry out security analysis tasks on USB file transfer used with the dashboards.

Participants in group C (security analysts), found the risk factors covered in the dashboard to be comprehensive and insightful, suggesting that some of the design results could also be applied to other internal security problems, such as screenshots and VPN access activity monitoring. Although the perceived usefulness and ease of use were high, the intention to use in the future was relatively low (B = 4.67). Since the scope of the dashboard was extremely narrow, a more integrated system that could monitor would be needed before some stakeholders would adopt the technology into their workflow.

Participants in groups B and (team managers and high-level security leaders) also noted that a wider range of security problems should be addressed before they adopt the system. They acknowledged the effectiveness of anomaly detection using visualization strategies, especially for users with limited technical and data analysis skills ($(U_{1(GroupA)}) = 7$, $N_{1(GroupA)} = 7$). The perceived usefulness for them in improving their productivity and reducing their workload was lower than for the security analysts since security monitoring is not one of their main work roles (C_1 = [3,5.5,5],C_2 = [3,4.5,5]). Thus, more sophisticated alerting and flagging systems would need to be developed to support a wider range of investigators in the detection process.

7.1. Usage report

In addition to the usability and intention to use evaluation above, observational data on the initial use of our dashboards in the organization are shown in Fig. 10. Usage numbers are shown by the department and by month in Figs. 10 and 11 (with detailed usage statistics being shown in Appendix D). Fig. 10 shows the usage count of the dashboard for each relevant department of the company in 2021. As presented in this figure, the security analytic team accessed the dashboard most frequently, as it was designed to support their day-to-day job functions.

Fig. 11 demonstrates the monthly usage count of the dashboards. In the first 2 quarters, stakeholders visited frequently to inspect anomalous USB file transfer activities. In the latter part of 2021, a strict control policy was implemented across the company to restrict all USB device uses at any endpoint, which then yield a massive drop in usage of the dashboard.

While the usage declined in later 2021, stakeholders at the company were aware of the excessive number of anomalies detected and

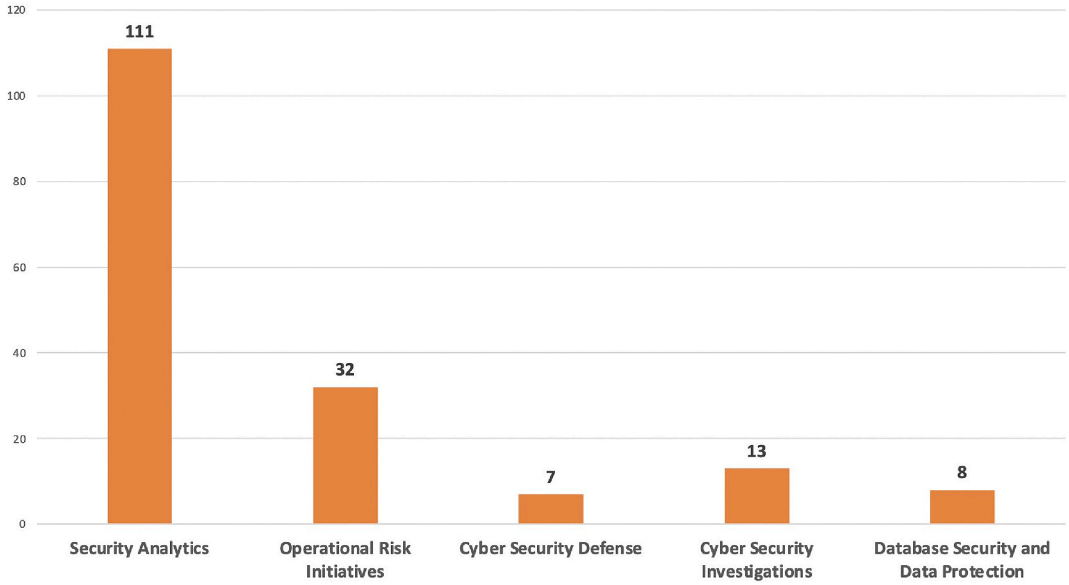


Fig. 10. Dashboard usage count by department.

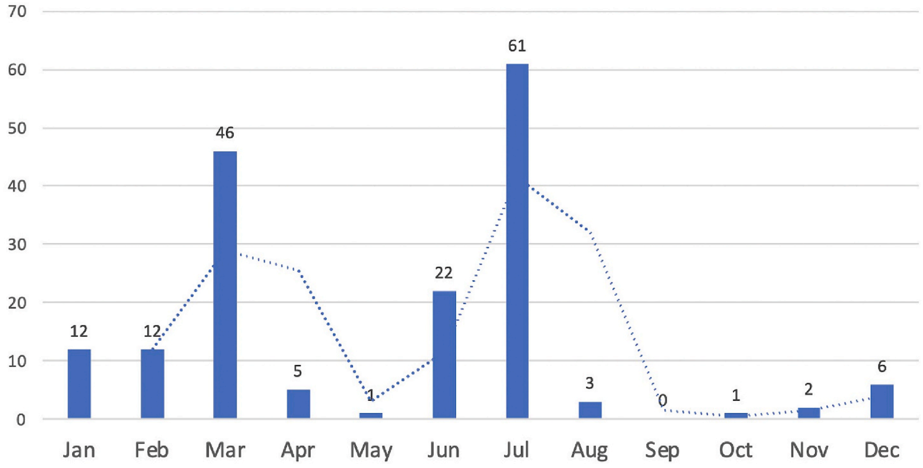


Fig. 11. Dashboard usage count and trendline by month.

visualized, which directly facilitated a policy update that USB sticks could no longer be used in the organization, thereby preventing potential threats. In addition, the concept and design lessons learned from building this dashboard tool led to similar dashboards being developed to monitor printing activities, email activities, VPN activities, DLP monitoring, etc. In general, the functionalities, interactivity, and learnability were well-received by the target user group and have improved the current detection strategies used by the company.

8. Discussion and limitation

Prior to the availability of our designed dashboard tool, alerts generated by cyber security experts or automated tools often contained a high rate of false positives. One of the reasons for this high rate was insufficient participation by all stakeholders in the preliminary investigation process. Thus in this study dashboards were developed which accommodate non-technical staff, improving their awareness of potentially malicious events, and making it easier to bring implicit knowledge to bear in detecting anomalies.

Another driving factor for using visualizations is that traditional alerts are incident-based. Expertise is thus required to form hypotheses and to conduct investigations, to aggregate multiple potentially relevant incidents together so as to detect a sequence of malicious events associated with a concerted exfiltration campaign. The new dashboard enabled high-level analysis based on file content and individual behavior during a selected period, reducing the number of alerts (sent to other investigators) and the likelihood of reporting false positives.

In the present study we focused on usage, and intention to use as the key outcome measures and we were driven by an urgent need to improve processes within the company. Due to these pressures, we did not measure the usability of the dashboards, except through the use of the TAM questionnaire, and we did not carry out finer-grained evaluations of the effectiveness of the dashboards. One consequence of this study was that the dashboards allowed the company to get detailed feedback (for the first time) on how USB sticks were being used. Management was surprised to find the extent to which the use of the USB sticks was potentially questionable and was generating anomalous behaviors. This led to two major impacts of the research reported here. First, the company hired one of the investigators (a graduating student) to build more dashboards that could be used to visualize its operations. Second, the company banned the use of USB sticks. Thus, it was not possible to carry out further studies on USB data exfiltration and it may be difficult to carry out similar studies with other companies, due to increased awareness about the risks that the use of USB sticks poses for large organizations. The present research is “a victim of its own success” and increased awareness of the risks associated with USB data transfers is likely restricting opportunities for future research on visualizing USB data in large organizations. However, the strategy used here for developing the dashboards, and the types of dashboards developed, can be generalized to other modes of data exfiltration.

9. Conclusion and future work

The research reported in this paper shows how visualization can be used to support the detection of possible data exfiltration anomalies in an industry setting. In this particular case, we developed a set of dashboards and filters to support the review of USB file transfer activity in a large financial services company. This work is important because proposed innovations in cybersecurity research need to be evaluated in practical settings. In this study not only did we construct visualization and filtering tools and collect user comments and suggestions for improvement, but we also used the technology acceptance model (TAM) to quantify the willingness of different types of users to use the novel visualization tools going forward.

The proposed visualization system addressed the needs of a wide range of stakeholders through a co-design approach. The dashboards we designed used integrated data visualization and interaction design, to permit nontechnical users to interpret otherwise overwhelming volumes of security data (collected through monitoring tools) without needing help from scarce security analyst resources. The early participation of all user groups in the design and preliminary investigation process improved detection performance, created better situational awareness, and produced lower error rates and shorter investigation times. Thus, visualization dashboard tools can provide actionable insights to a wide range of investigators and the dashboards designed here represent a first step toward the development of more intelligent and integrated detection tools.

Dashboard design should be just the first step in a more comprehensive defense against data exfiltration that is also AI-driven. Visualization tools can be used to efficiently label malicious events using an Interactive Machine Learning (iML) approach [39]. Visualization and iML can be combined to improve the efficiency of training and the ultimate accuracy of anomaly detection models. Another recommended strategy for use with visualization dashboards is to use computer monitoring tools on this platform to collect information on the work behavior of different user groups, develop comprehensive task analyses and improve the performance and associated interactive machine learning algorithms [40].

Ethics approval statement

The interviews, focus group, and required analysis procedures were approved by the University of Toronto ethics review board (Ethics protocol number 39752).

Author contribution statement

Mu-Huan Chung: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Yuhong Yang: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Lu Wang; Greg Cento; Khilan Jerath; Parwinder Taank; Abhay Raman; Jonathan Hoyin Chan: Contributed reagents, materials, analysis tools or data.

Mark Chignell: Conceived and designed the experiments; Wrote the paper.

Funding statement

Mark Chignell was supported by Mitacs [512048].

Data availability statement

The data that has been used is confidential.

Declaration of interest's statement

The authors declare the following conflict of interests: One of the authors is the section head for Computer Science at Heliyon.

Appendix A. Assumption and rationale of each filter designed for the dashboards

#	Filter	Assumption and Rationale
1	Employee Name/ Employee ID	Since employees would typically behave in keeping with their historical records, the filter should enable individual-based aggregation and comparison
2	Business Group	Employees should behave similarly with colleagues in the same department; thus, the filter should enable division-based aggregation and comparison
3	Job Role	Employees should behave similarly with colleagues in the same role; thus, the filter should enable role-based aggregation and comparison
4	Manager	Employees should behave similarly with colleagues in the same team; thus, the filter should enable team-based aggregation and comparison
5	Employee Type	USB transfers carried out by non-regular employees (i.e., temporary, contractor, fixed-term, ven- dor, casual employees) are assumed more likely to be malicious than transfers by regular/full-time employees.
6	Employee Status	USB file transfers by inactive employees are assumed to be potentially malicious.
7	Admin Access	Admin users have access to read, write, and transfer critical corporate information as part of their job, which needs to be monitored carefully.
8	USB Exemption	USB transfers done by employees without authorized access are potentially malicious.
9	GCE Access	Abnormal file transfers done by employees with governmental-related data access are potentially more damaging to the company and should be monitored carefully.
10	High-Risk File Type	File transfer of high-risk file types, such as data files, zip files, are assumed to be riskier than other file types, such as.mp4 or. temp files.
11	Outside Working Hours	USB File transfers outside normal business hours are of concern, unless the transfers are done by a person who is working in a different time zones with different business hours.
12	Termination Date	Files extracted by employees close to their termination date are worth investigating to avoid infor- mation leaks due to the employee being disgruntled or opportunistic.

Appendix B. Task-based usability testing protocol

Security analysts

Analysts are the main users of the dashboard, who regularly monitor file copying events, give suggestions, and generate potential anomaly alerts, export data visualization reports for managers and investigators. The goal of the testing is to verify whether the analysts can use the dashboard to complete their **anomaly detection and monitoring tasks** and **generate real-time reports** for the other investigators.

Part 1: Ask the participant to complete ***two of the analytics tasks*** and ***one exploratory task*** and ***generate a report*** to a manager with specific requirements.

●How many employees transferred files to their own device in your own team?
●How many files have been transferred to a personal USB device in the last week in a department?
●Identify top 3 employees with the highest amount of USB file transfers.
●Search who transferred files related to keywords “xxx”?
●Identify employees who transferred files to their personal devices outside working hours.
●Identify employees who transferred files from more than one computer.
●Identify employees who transferred files to more than one USB Device.
●Explore the USB file transfer activity of a specific employee that you know.
●Explore the USB file transfer activities of employees near the termination date.
●Export the data visualization report for your manager (under special requirements).

Part 2: SUS + Survey question.
Answer the following question after the usability testing.

●Does the dashboard so far provide all the information you need on the topic? If not, what is missing?
●Was the navigation smooth and clear? If not, why?
F0B7 Do figures seem correct to you? If not, please list the potential errors you have spotted.
●Does the tool help you to explain the reason for malicious to other roles?
●Does the tool help you to generate informative report to other roles, such as to department manager? Complete the following System Usability Scale (SUS)

This is a standard questionnaire that measures the overall usability of a system. Please select the answer that best expresses how you feel about each statement after using the website today.

Other Stakeholders

The **team leader, file owner, high-level administrators and other investigators** will receive the regular activity reports to monitor the USB Activities, or validate any anomalous incidents identified by security analysts and use the dashboard to further investigate the individual if needed. The goal of the testing is whether they can **understand the static report** and whether they can **complete explor-atory analysis** on their own.

Part 1: Ask the participant to ***answer 2 questions from looking at a static report*** and complete ***one exploratory task***:
Answer the following question by looking at the report.

●Who transferred the most files last week?
F0B7 How many employees transferred files to their own device in your own team?
F0B7 How many files have been transferred to a personal USB device in the last week in a department?
F0B7 Identify employees who transferred files to their personal devices outside working hours.
F0B7 Identify employees who transferred files from more than one computer.
F0B7 Identify employees who transferred files to more than one USB Device.

Exploratory Tasks

●Explore the USB file transfer activity of a specific employee that security analysts identified as malicious.
●Explore your team’s USB File activities in the XXX project timeline.

Part 2: In addition to the general survey questions and SUS assessment, we should also verify whether the other stakeholders can un-derstand the static report generated from the dashboard.

●Does the report so far provide all the information you need on the topic? If not, what is missing?
●Can you understand why this individual was identified as malicious from the static report? If not, which part is confusing?

Appendix C

Adjusted Technology Acceptance Model Questionnaire

Appendix D. Dashboards Usage Report

Month	Department - User	Count
January	Security Analytics 4	12
February	Security Analytics 1	4
	Security Analytics 4	8
March	Security Analytics 1	12
	Security Analytics 4	7
	Operational Risk Initiatives 1	23
	Operational Risk Initiatives 2	4
April	Operational Risk Initiatives 2	5
May	Security Analytics 4	1
June	Security Analytics 1	6
	Security Analytics 3	8
	Cyber Security Defense 1	2
	Database Security and Data Protection 1	6
July	Cyber Security Investigations 1	10
	Security Analytics 2	39
	Security Analytics 4	5
	Cyber Security Defense 1	5
	Database Security and Data Protection 1	2
August	Cyber Security Investigations 1	3
September	N/A	0
October	Security Analytics 1	1
November	Security Analytics 2	2
December	Security Analytics 2	6

References

[1] I. Homoliak, F. Toffalini, J. Guarnizo, Y. Elovici, M. Ochoa, May, Insight into insiders and IT: a survey of insider threat taxonomies, analysis, modeling, and countermeasures, ACM Comput. Surv. 52 (2) (2019) 40. Article 30.

[2] S.M. Bellovin, The Insider Attack Problem Nature and Scope" in Insider Attack and Cyber Security, 2008, pp. 1–4.

[3] B. Sabir, F. Ullah, M.A. Babar, R. Gaire, Machine learning for detecting data exfiltration: a review, ACM Comput. Surv. 54 (3) (2021) 1–47.

[4] M. Evangelou, N.M. Adams, Oct, An anomaly detection framework for cyber-security data, Comput. Secur. 97 (2020).

[5] M. Whitehouse, M. Evangelou, N.M. Adams, Activity-based temporal anomaly detection in enterprise cyber security, in: Proceedings of 2016 IEEE Conference on Intelligence and Security Informatics: ISI, 2016, pp. 248–250.

[6] M. Ahmed, A.N. Mahmood, R. Islam, A survey of anomaly detection techniques in financial domain, in: Future Generation Computer Systems 55, 2016, pp. 278–288.

[7] F. Ullah, M. Edwards, R. Ramdhany, R. Chitchyan, M.A. Babar, A. Rashid, Data exfiltration: a review of external attack vectors and countermeasures, J. Netw. Comput. Appl. 101 (2018) 18–54.

[8] A. Khraisat, I. Gondal, P. Vamplew, Survey of intrusion detection systems: techniques, datasets and challenges, Cybersecurity 2 (2019) 20.

[9] L. Jiang, A. Jayatilaka, M. Nasim, M. Grobler, M. Zahedi, M.A. Babar, Systematic Literature Review on Cyber Situational Awareness Visualizations, IEEE Access, 2022.

[10] C. Gates, S. Engle, Reflecting on visualization for cyber security, in: Proceedings of 2013 IEEE Inter-national Conference on Intelligence and Security Informatics, 2013, pp. 275–277.

[11] D. Staheli, V. Mancuso, R. Harnasch, C. Fulcher, M. Chmielinski, A. Kearns, S. Kelly, E. Vuksani, Collaborative data analysis and discovery for cyber security, in: Proceedings of the Twelfth Symposium on Useable Privacy and Security: SOUPS '16, 2016.

[12] M.R. Endsley, Design and evaluation for situation awareness enhancement, in: Proceedings of the Human Factors Society Annual Meeting 32, SAGE Publications, Sage CA: Los Angeles, CA, 1988, pp. 97–101.

[13] E.D. Matthews, H.J. Arata III, B.L. Hale, Cyber situational awareness, Cyber Def. Rev. 1 (1) (2016) 35–46.

[14] P. Barford, M. Dacier, T.G. Dietterich, M. Fredrikson, J. Giffin, S. Jajodia, X. Ou, Cyber SA: situational awareness for cyber defence, in: Cyber Situational Awareness, Springer, Boston, MA, 2010, pp. 3–13.

[15] D. Napoli, Developing accessible and useable security (ACCUS) heuristics, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems: CHI'18, 2018, pp. 1–6.

[16] A. Whitten, J.D. Tygar, Usability of Security: A Case Study, Carnegie Mellon University, 1998.

[17] J. Rode, C. Johansson, P. DiGioia, R.S. Filho, K. Nies, D.F. Nguyen, J. Ren, P. Dourish, D. Redmiles, Seeing further: extending visualization as a basis for useable security, in: Proceedings of the Second Symposium on Useable Privacy and Security: SOUPS '06 145–155, 2006.

[18] C. Carreira, J.F. Ferreira, A. Mendes, N. Christin, Exploring Useable Security to Improve the Impact of Formal Verification: a Research Agenda, 2021 arXiv preprint arXiv:2111.08209.

[19] D. Staheli, T. Yu, R. Jordan Crouser, S. Damodaran, K. Nam, D. O'Gwynn, S. McKenna, L. Harrison, Visualization evaluation for cyber security: trends and future directions, in: Proceedings of in the Eleventh Workshop on Visualization for Cyber Security: VizSec '14, 2014.

[20] S. Parkin, A. Moorsel, P. Inglesant, M.A. Sasse, A stealth approach to useable security: helping IT security managers to identify workable security solutions, in: Proceedings of the 2010 New Security Paradigms Workshop: NSPW '10, 2010, pp. 33–50.

[21] S.L. Pfleeger, J.B. Predd, J. Hunker, C. Bulford, March, Insiders behaving badly: addressing bad actors and their actions, IEEE Trans. Inf. Forensics Secur. 5 (1) (2010) 169–179.

[22] M.L. Griffin, M.P. Harman, R.G. Shiflet Jr., T.H. Stigler, D.G. Turner, D.D. Turner, Comprehensive suspicious activity monitoring and alert system, in: United States Patent US 8.412,605 B2, 2 4, 2013.

[23] I. Lee, Internet of Things (IoT) cybersecurity: literature review and IoT cyber risk management, Future Internet 12 (9) (2020) 157.

[24] B. Haim, E. Menahem, Y. Wolfsthal, C. Meenan, Visualizing insider threats: an effective interface for security analytics, in: Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion: IUI '17 Companion, 2017, pp. 39–42.

[25] S. Chen, C.P. Janeja, Human perspective to anomaly detection for cybersecurity, J. Intell. Inf. Syst. 42 (1) (2014) 133–153.

[26] I.A. Gheyas, A.E. Abdallah, Detection and prediction of insider threats to cyber security: a systematic literature review and meta-analysis, Big Data Anal 1 (2016) 6.

[27] A. Gamachchi, L. Sun, S. Boztas, A graph based framework for malicious insider threat detection, in: Proceedings of 50th Hawaii International Conference on System Sciences: HICSS 50, 2017.

[28] P.A. Legg, Visualizing the insider threat: challenges and tools for identifying malicious user activity, in: Proceedings of 2015 IEEE Symposium on Visualization for Cyber Security: VizSec, 2015, pp. 1–7.

[29] J.J. Hanniel, T.E. Widagdo, Y.D.W. Asnar, Information system log visualization to monitor anomalous user activity based on time, in: Proceedings Pf 2014 International Conference on Data and Software Engineering: ICODSE'14, 2014, pp. 1–6.

[30] M.H. Chignell, M.-H. Chung, Y. Yang, G. Cento, A. Raman, Human factors in interactive machine learning: a cybersecurity case study, Proc. Hum. Factors Ergon. Soc. Annu. Meet. 65 (1) (2021) 1495–1499.

[31] S.Y. Park, An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning, Educ. Technol. Soc. 12 (3) (2009) 150–162.

[32] V. Venkatesh, F. Davis, A theoretical extension of the technology acceptance model: four longitudinal field studies, Manag. Sci. 46 (2000) 2.

[33] Y. Lee, K.A. Kozar, K.R.T. Larsen, The technology acceptance model: past, present, and future, in: Proceedings of Communications of the Association for Information Systems 12, 2003, p. 50.

[34] F. Davis, R. Bagozzi, P. Warshaw, User acceptance of computer technology: a comparison of two theoretical models, Manag. Sci. 35 (1989) 8.

[35] S.A. Kamal, M. Shafiq, P. Kakria, Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM), Technol. Soc. 60 (2020), 101212.

[36] S.A. Salloum, A.Q.M. Alhamad, M. Al-Emran, A.A. Monem, K. Shaalan, Exploring students' acceptance of e-learning through the development of a comprehensive technology acceptance model, IEEE Access 7 (2019) 128445–128462.

[37] Z. Hu, S. Ding, S. Li, L. Chen, S. Yang, Adoption intention of fintech services for bank users: an empirical examination with an extended technology acceptance model, Symmetry 11 (3) (2019) 340.

[38] A.M. Mutahar, N.M. Daud, T. Ramayah, O. Isaac, A.H. Aldholay, The effect of awareness and perceived risk on the technology acceptance model (TAM): mobile banking in Yemen, Int. J. Serv. Stand. 12 (2) (2018) 180–204.

[39] S. Amershi, M. Cakmak, W.B. Knox, T. Kulesza, Power to the people: the role of humans in interactive machine learning, AI Mag. 35 (4) (2014) 105–120.

[40] C. Mu-Huan, C. Mark, W. Lu, J. Alexandra, R. Abhay, Interactive machine learning for data exfiltration detection: active learning with human expertise, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020, October, pp. 280–287. IEEE.