UiA University of Agder

Master

Designing a framework for data populating alarms based on MITRE techniques

How can the MITRE ATT&CK framework be utilized to automate information gathering for security events in a SOC environment?

SVERRE OSE DYBDAHL and MARTIN NAUF STAER

SUPERVISORS
Nadia Saad Noori & Erlend Halsnes

University of Agder, 2023
Faculty of Engineering and Science
Department of Engineering and Sciences

Obligatorisk gruppeerklæring

Den enkelte student er selv ansvarlig for å sette seg inn i hva som er lovlige hjelpemidler, retningslinjer for bruk av disse og regler om kildebruk. Erklæringen skal bevisstgjøre studentene på deres ansvar og hvilke konsekvenser fusk kan medføre. Manglende erklæring fritar ikke studentene fra sitt ansvar.

1.	Vi erklærer herved at vår besvarelse er vårt eget arbeid, og at vi ikke har brukt andre kilder eller har mottatt annen hjelp enn det som er nevnt i besvarelsen.	Ja
2.	**Vi erklærer videre at denne besvarelsen:** • Ikke har vært brukt til annen eksamen ved annen avdeling/universitet/høgskole innenlands eller utenlands. • Ikke refererer til andres arbeid uten at det er oppgitt. • Ikke refererer til eget tidligere arbeid uten at det er oppgitt. • Har alle referansene oppgitt i litteraturlisten. • Ikke er en kopi, duplikat eller avskrift av andres arbeid eller besvarelse.	Ja
3.	Vi er kjent med at brudd på ovennevnte er å betrakte som fusk og kan medføre annullering av eksamen og utestengelse fra universiteter og høgskoler i Norge, jf. Universitets- og høgskoleloven §§4-7 og 4-8 og Forskrift om eksamen §§ 31.	Ja
4.	Vi er kjent med at alle innleverte oppgaver kan bli plagiatkontrollert.	Ja
5.	Vi er kjent med at Universitetet i Agder vil behandle alle saker hvor det forligger mistanke om fusk etter høgskolens retningslinjer for behandling av saker om fusk.	Ja
6.	Vi har satt oss inn i regler og retningslinjer i bruk av kilder og referanser på biblioteket sine nettsider.	Ja
7.	Vi har i flertall blitt enige om at innsatsen innad i gruppen er merkbart forskjellig og ønsker dermed å vurderes individuelt. Ordinært vurderes alle deltakere i prosjektet samlet.	Nei

Publiseringsavtale

Fullmakt til elektronisk publisering av oppgaven Forfatter(ne) har opphavsrett til oppgaven. Det betyr blant annet enerett til å gjøre verket tilgjengelig for allmennheten (Åndsverkloven. §2). Oppgaver som er unntatt offentlighet eller taushetsbelagt/konfidensiell vil ikke bli publisert.

Vi gir herved Universitetet i Agder en vederlagsfri rett til å gjøre oppgaven tilgjengelig for elektronisk publisering:	Ja
Er oppgaven båndlagt (konfidensiell)?	Nei
Er oppgaven unntatt offentlighet?	Nei

Acknowledgements

First of all we would like to thank Netsecurity for giving us the opportunity to work with them for this project. We would also like to thank our supervisors Nadia Noori from the University of Agder (UiA) and Erlend Halsnes from Netsecurity, whom both have been supportive and helpful along the way, whenever we have had any problems or uncertainties. In hindsight perhaps we should have initiated a tighter cooperative relationship with both during the project.

Abstract

In this paper we aim to develop a proof of concept framework as a step-by-step process for identifying what type of information and log types a SOC analyst needs to analyze and handle an alarm based on the alarms MITRE technique. To solve this, it was decided that using both theoretical and experimental research methodologies could be advantageous. Hence we first used a Systematic Literature Review to search, screen, and select relevant literature. Followed by the usage of Design Science Research method for conducting the research based upon a theoretical basis, and an experimental process. To develop a framework consisting of an easy to understand and independent step-by-step process.

The proof of concept framework introduced in this paper, is an eight step process describing how one may proceed when gathering data needed for automating information gathering based on alarms MITRE techniques. In these eight steps it revolves around three main concepts, which are gathering a theoretical foundation by research and discussion, improving the theoretical foundation by testing and adjusting, and ends with a continuous process of maintaining the constructed automations when used in a production setting. This framework produced accurate results when tested during research, and we believe it should be further explored and tested in a larger scale. Also it should be considered a stepping stone into further automating the whole alarm handling process, from gathering data to response.

Contents

List of Figures

List of Tables

Abbreviations	
AI	Artificial Intelligence
APT	Advanced Persistent Threat
ATT&CK	Adversarial Tactics, Techniques and Common Knowledge
CSOC	Cyber Security Operation Center
DSR	Design Science Research
ICT	Information and Communication Technology
IoT	Internet of Things
IT	Information Technology
MS	Microsoft
MSSP	Managed Security Service Provider
SEM	Security Event Management
SIM	Security Information Management
SIEM	Security Information and Event Management
SOAR	Security Orchestration, Automation and Response
SOC	Security Operation Center
SLR	Systematic Literature Review
STIX	Structured Threat Information eXpression
TAXII	Trusted Automated eXchange of Intelligence Information

Chapter 1

Introduction

Netsecurity is a Managed Security Service Provider(MSSP) who is planning to make a change in their alarm handling to base it on the MITRE ATT&CK framework instead of having it based on the specific customers. If this is possible, it would enable the possibility of generalizing the alarm handling more and therefore make it possible to scale up their operation more rapidly in terms of recruiting new customers by having less customer specific automations.

Therefore, Netsecurity asked us to look into this as a third part to look at the case with new eyes and perspectives. At first we were considering if we were going to base our solution on the MITRE tactics, or the MITRE techniques. Where we opted to base our solution on the techniques, to find out what information a security analyst would need to be able to analyze and handle an alarm.

Due to the amount of MITRE techniques, the scope of the task and that we were working on a restricted timeline, we came to the conclusion that the most optimal for both parts would be that we would focus on creating a framework to figure out if it is possible to identify what type of data that is required to make a verdict for an alarm based on a specific technique, whether or not it would be feasible. And how the process of identifying the required data for a technique could look like. From there came the research question *"How can the MITRE ATT&CK framework be utilized to automate information gathering for security events in a SOC environment?"*

1.1 Limitations

During the project we had some limitations, both foreseen limitations and unforeseen ones, which we had to adjust by. The foreseen limitations where the limited time as previously mentioned for the project and that we had limited prior knowledge about working with playbooks and its complexity. If we had this prior knowledge we would have been able to create playbook automations for testing, which would most likely would give us the ability to test more techniques and more alarms per technique, which could have resulted in a more accurate result. While the unforeseen limitations was that we were unable to find any relevant equivalent research on the matter which we could base our research upon. Another unforeseen limitation was that things took longer time than expected such as for example the SLR process, giving us less time to focus on the research itself.

1.2 Thesis Structure

- **Chapter 1. Introduction :** Provides a brief explanation of how the thesis came to life, the goal of the thesis and the limitations. Before it rounds of with this overview of the thesis structure.

- **Chapter 2. Literature Review and Theory :** Introduces the SLR method and how we have we have used it for gathering literature for the thesis, before briefly explaining theory of relevant topics found from the SLR.

- **Chapter 3. Methodology :** The decision was made to use DSR for this research. In this chapter we explain what DSR is, why we chose it, and how we utilized it in our research.

- **Chapter 4. Result - A Step by Step Framework :** Presents the step by step framework we have come up with and goes through each step of the framework.

- **Chapter 5. Discussion :** Is where we summarize our findings and explain the interpretations of the results, before we go through the limitations and end by sharing our recommendations for the framework.
- **Chapter 6. Conclusion :** In the conclusion chapter we offer our final notes in how we believe the research question and other requirements were answered, we also explain how this research contributes to filling a gap in the existing literature, while simultaneously wrapping up this paper.

Chapter 2

Literature Review and Theory

This chapter will first define Systematic Literature Review (SLR) and the process of how it has been utilized for gathering research theory for this paper. A complete list of included theory is also present. Thereafter focus on defining relevant topics of theory such as MITRE, SOAR etc. based on theory gathered during the SLR process.

2.1 Literature Review

In order to progress our study we need to understand the fundamentals around frameworks, methodologies and concepts that the research question encompasses. Furthermore, it is critical when performing research to identify what work has been performed on the same or similar topics, and potentially use previous work as a foundation, or use the lack of it as motivation. This problem can be solved by conducting a Systematic Literature Review (SLR), which can be defined as such: *"a means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest"*. [5] Our rationale for choosing SLR over other literature review methodologies is the fact that it offers an extensive, replicable, and qualitative process which would be necessary given the scarce amount of similar research, as well as our intention to make this thesis a building block for future work.

2.1.1 Methodology

We elected to use a Systematic Literature Review (SLR) to build a basis for our research, following a methodology as depicted in figure 2.1. SLR can be advantageous compared to other literature methodologies because it can offer a higher degree of quality, replicability, reliability, and validity of reviews. [10] Further, for any novice researchers such as ourselves, the ability to reproduce results could be critical for supervisors and guidance counselors to aid our research. Also the iterative process of it all proves invaluable as trial and error is a key part in identifying usable search terms to cover more relevant research.

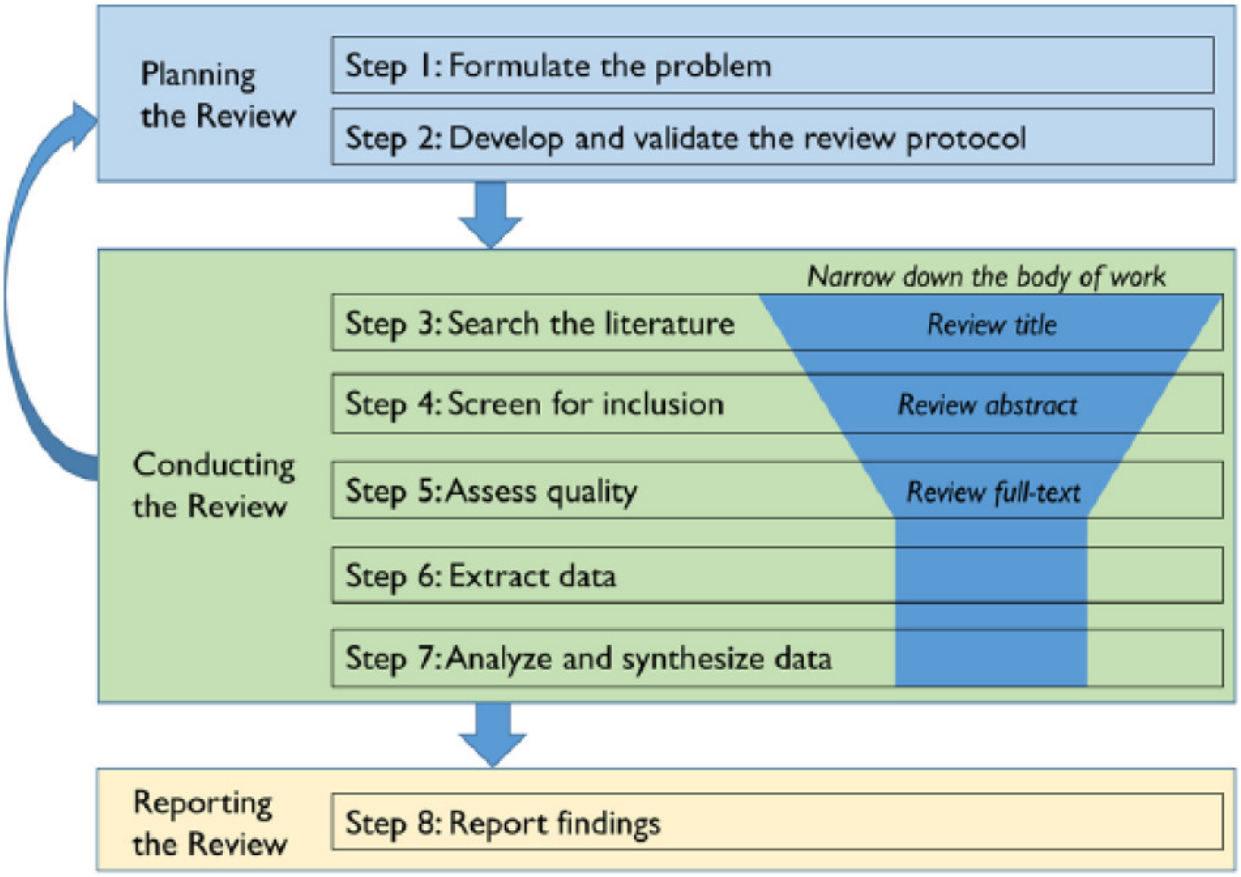

Figure 2.1: SLR Methodology [10]

2.1.2 Planning Process

Based on the works of Xiao and Watson [10], we can identify the second step of the SLR as the protocol development and validation stage. And they further state that this protocol "*is a preset plan that specifies the methods utilized in*

conducting the review". For our thesis, we were experiencing time constraints and elected to focus on three elements of this process in general.

Inclusion Criteria

- Must be written in Norwegian or English.
- Must be relevant for our keywords and theme.
- Must be written in the last 10 years, and newer publications would be prioritized.
- Must be peer reviewed.
- Must be published in a recognized journal.

During our search process we came across some texts that were not to be considered an academic text, and thus not necessarily peer reviewed. However we elected to make an exception for those texts given their relevance to our thesis. Careful considerations were made when selecting material not discovered in the literature review, in order to maintain a level of professionalism regarding the source material.

Search Strategies

To begin the search process we had to identify some relevant sources to gather information. Xiao and Watson identifies three major sources for information: electronic databases, backward search and forward search.[10] We elected to also use this methodology and decided on the following electronic databases:

- IEEE Xplore
- Scopus
- Google Scholar
- Web of Science

In order to figure out what some relevant searches for our thesis might be, we made a few keywords based on our primary objective for this thesis(see table 2.1 below). And from those keywords we identified synonyms and similar words that could be used in phrasing the search queries.

Keywords	**Synonyms / Similar Words**
Cybersecurity	Cyber security, IT security, Information Security
MITRE	MITRE matrix, MITRE ATT&CK, ATT&CK, MITRE attack, MITRE framework, MITRE tactics
Security Operation Center	SOC, SOC Efficiency
Automated Alarm Analysis	SOAR, Security Orchestration, Automation and Response, Triage
Security Information and Event Management	SIEM

Table 2.1: keywords

At this stage we started putting together an initial list of 20 phrases based on those keywords and their synonyms and similar words. We chose not to use them all in combination as that would create too large of an amount of searches for us to perform in the next stage of the project, which would take too much time away from the research phase.

1. "SOC" and "Cybersecurity"
2. "SOC" and "Information Security"
3. "SOC" and "IT Security"
4. "SOC Efficiency" and "MITRE"
5. "SOC Efficiency" and "Automated Alarm Analysis"

6. "SOC Efficiency" and "SOAR"
7. "SOC" and "Automated Alarm Analysis"
8. "Cybersecurity" and "Automated Alarm Analysis"
9. "Security Operations Center"
10. "MITRE" and "SOC"
11. "ATT&CK" and "SOC"
12. "MITRE" and "SOAR"
13. "ATT&CK" and "SOAR"
14. "MITRE" and "SIEM"
15. "ATT&CK" and "SIEM"
16. "MITRE" and "Triage"
17. "ATT&CK" and "Triage"
18. "MITRE" and "Automated Alarm Analysis"
19. "ATT&CK" and "Automated Alarm Analysis"
20. "SIEM" and "Automated Alarm Analysis"

What remained at this stage would be to determine two things. Seeing as we are two students working on this thesis, would it be optimal to split the databases so that we would do two each, or do them all and compare results? Also, where should we set the limit as to how many search results we look through? Seeing as for some queries, big databases such as Google Scholar would return tens of thousands of results.

We elected to both search through all the databases, reasoning being that we would be very likely to identify at least a few articles that the other part would not. It is written in the works of Kitchenham and Charter [5] that it is advantageous to cooperate during the inclusion process in order to maintain validity. This would create more work, but the benefits would outweigh this issue. For the second problem we did some test searches and discovered that for the most part, we would find the articles that looked the most interesting in the first ≈ 100 results, as such we set a limit to the first 150 results to try to cover the most ground while not reading through thousands of headlines per query.

Screening procedures

We identified that the main part of our screening process would consist of three parts. Searching for literature would be first, in which we would perform the search and select articles based on their title. Following that would be a phase where we would read the abstract of the articles collected in phase one, and select those with relevant abstracts. Then using the results from phase two, we would read the full articles to determine if they were suited for our research project. This methodology aligns with the procedure created by Xiao and Watson [10] as depicted in figure 2.1

Following this procedure we also identified that we would perform a backward and forward search based on our findings, and put them through the same three step process as above. Furthermore we planned that it would be likely that we'd need to supplement the search with more phrases if we discovered some shortcomings in the process this is in accordance to Xiao and Watson [10] who state that "*Deeper understanding can be gained during the review process, requiring a change in keywords and/or analytical methods*", those shortcomings would also follow the same three step process.

2.1.3 Search and Screening Process

Following the plan, we both went ahead to perform the searches individually, collecting all relevant papers based on their title in our respective Google Sheets documents, keeping the title, search phrase, date and database. Following this process we gathered all the results in a singular document and removed duplicates, and copy if this document is available in appendix A. Reviewing abstracts were done by reading them and discussing one by one if they seemed relevant to our thesis. The full text reviews were then done by reading each of the remaining papers, and making notes of relevancy within the papers to compare and discuss which ones to keep. The results of each phase of this process can be observed in figure 2.2. Following this, we performed the backward and forward search in the same manner, as depicted in figure 2.3, although we did not select any of the articles we found in that process.

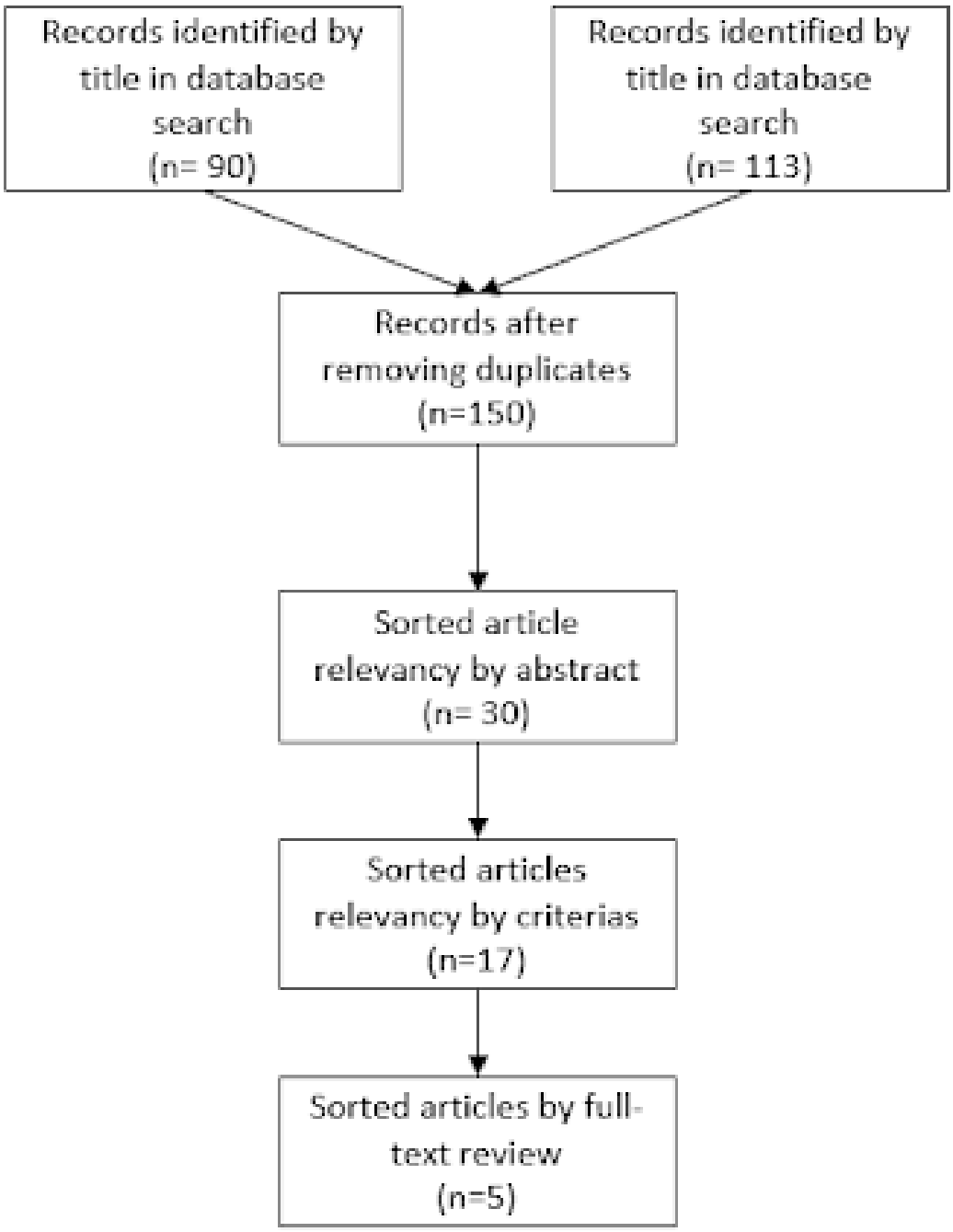

Figure 2.2: Main Search Results

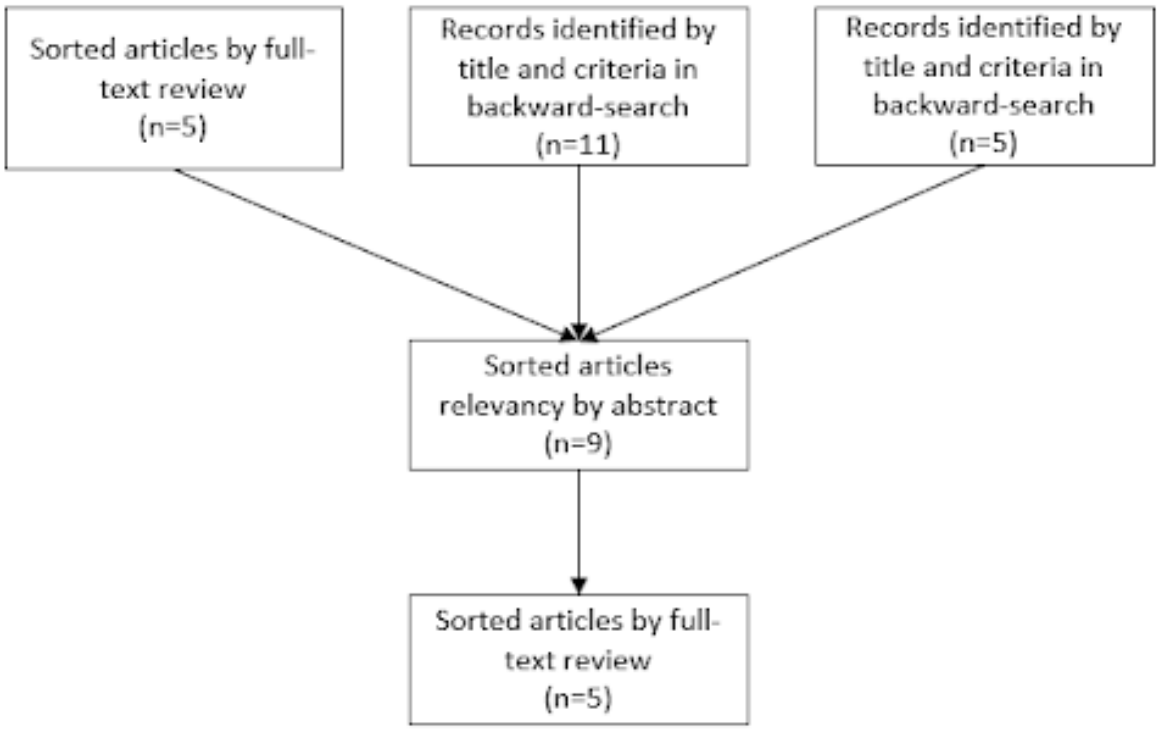

Figure 2.3: Backward and Forward Search Results

Supplementing the searches in later time is not a bad idea and should be expected in a SLR, this was also the case for our thesis, in which we included some extra search terms based on our supervisors feedback. [10] The results of this supplementary search can be seen in figure 2.4, while a detailed document can be found in appendix B. And the supplemented search terms are as following:

- Cyber Threat Intelligence
- Cyber Forensics
- Cyber Incident Triage Process
- Cyber Incident Triage Systems
- Automation of Cyber Incident Triage
- Bayesian Network Cyber Incident Triage

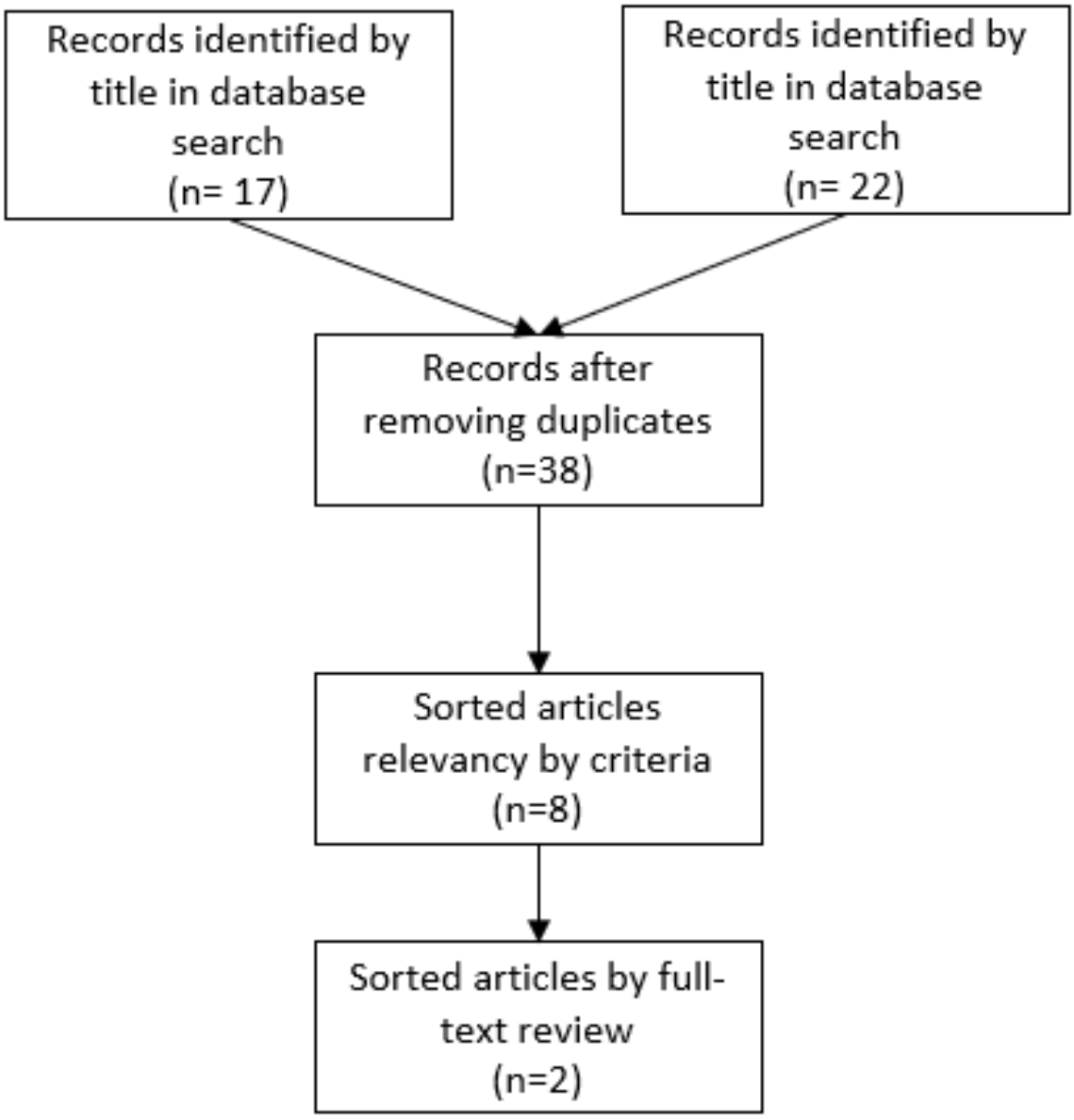

Figure 2.4: Supplementary Search Results

2.1.4 Results

Following the search and screening phase, we were left with 7 sources we thought relevant to our thesis as displayed in the table 2.2 below. One can argue that this is less than what may be expected of performing an SLR, but we reiterate the definition of SLR from Kitchenham and Charters [5], "*A means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest*" and remember that SLR is also an important tool to identify relevant source material or lack thereof. Such as our problem which is reasonably new, and not currently widely

applicable, which can therefore be considered as a gap in existing source material.

Title	**Author**	**Year**	**Source**	**Keywords**
A Framework for Designing a Security Operations Centre (SOC)	Stef Schinagl Keith Schoon, Ronald Paans	2015	IEEE Xplore	SOC, framework, cybercrime, IT Abuse, Value, sharing knowledge, secure service development, continuous monitoring, damage control, model, Intelligence baseline security, monitoring, pentest, forensic
An Overview of MITRE Cyber Situational Awareness Solutions	MITRE CORP MCLEAN VA MCLEAN	2015	Defence Technical Informa-tion Center	MITRE, STIX, TAXII, MITRE matrix, Crown jewel analysis, Cyber Command Units, Threat Assessment and Remediation Analysis
Assessing MITRE ATT&CK Risk Using a Cyber-Security Culture Framework	Anna Georgiadou, Spiros Mouzakitis, Dimitris Askounis	2021	Web Of Science	cyber-security culture framework, MITRE ATT&CK matrix, security assessment, detection, mitigation techniques

Improving SIEM alert metadata aggregation with a novel kill-chain based classification model	Blake D. Bryant, Hossein Saiedian	2020	Science Direct	Network monitoring, Intrusion detection, Kill-chain, Advanced persistent threat, APT, Security information and event management, SIEM, Security log ontology, Computer network defense, Attack ontology, Threat framework
Pushing the Limits in Event Normalisation to Improve Attack Detection in IDS/SIEM Systems	Amir Azodi David Jaeger, Feng Cheng, Christoph Meinel	2013	IEEE Xplore	Event Normalisation, Intrusion Detection, Event Management, Knowledge base
Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy	Cyril Onwubiko	2015	Google Scholar	Cyber Security Operations Centre, CSOC, SOC, Cyber Incident Response, Cyber Situational Awareness, CyberSA, Log Source, Analysis, Correlation, Risk Management, CSOC Strategy, CSOC Benefits & Challenges

Finding cyber threats with ATT&CK-based analytics	MITRE CORP ANNAPOLIS JUNCTION MD	2017	Google Scholar	Mitre, Tactics, Techniques, Endpoint Sensing, ATT&CK, ATT&CK-bases, analysis, Threat based security,

Table 2.2: SLR results

2.2 MITRE ATT&CK

MITRE ATT&CK is a framework made to describe the actions an adversary takes when attempting to, or while operating within an enterprise network. The framework was initiated in 2013 by the MITRE corporation, and the intention behind it was to *"document and categorize post-compromise adversary tactics techniques and procedures (TTPs) against Microsoft Windows systems aiming to improve detection of malicious behavior"*. [4] Since then, its objective has slightly changed, along with its content. For example, while it was originally aimed towards Windows systems, in recent time it also includes Linux systems. As well as its tactics manly entailed post-compromise adversary tactics, it now includes reconnaissance and resource development both of which may happen before any compromise. [11] MITRE ATT&CK consists in large part of a set of tactics and techniques used by Advanced Persistent Threats (APTs), collected by real reports of APT intrusions by the public, as well as mitigation methodologies toward these. [12] It also contains infromation on different APT groups, such as attack methodologies, tools and techniques used, contributed to by the public.

MITRE ATT&CK Components

Some important features of MITRE ATT&CK are as follows: tactics, techniques, mitigations and adversary groups. Where each of these categories are

largely contributed to based on public reports of real attacks.

- Tactics
  A tactic in MITRE ATT&CK is a collection of techniques, and it denotes to a degree the objectives of performing the actions of the techniques witin. [4][12] As of today there are 14 different tactics which contains methodologies from topics ranging from phases thoughout an attack, starting with reconnaissance and ending with impact.[11] As such it may seem like the tactics seem to follow other models of an adversaries life cycle, or rather a kill chain. However it is to be noted that in ATT&CK, the tactics are meant to denote the actions on a singular endpoint as an adversary moves through the network, whereas kill chains often show the broader scope that entails an entire operation. [12]

- Techniques
  Whereas tactics descibe the goal of an operation, techniques describe the operation being performed to reach the goal. [4] For each tactic there exists several techniques as methods to reach the goal of the tactics. For example for the tactic initial access there exists a technique called phishing, where an adversary can obtain access to an account by utilizing phishing. Furhtermore techniques contain detection and mitigation methodologies, and some contain sub-techniques, which in the case of phishing is different kinds of spear phishing. [11]

- Mitigations
  Each technique contain a subcategory that is mitigations, which exists to define countermeasures that could be taken to prevent an adversary in achieving their goal thorough the use of said technique. [4] Mitigations can range from methods such as user training, to access management and backups procedures [11]

- Adversary group
  Albeit somewhat detached from the rest of the elements described, MITRE ATT&CK also contains a database of adversary groups. Often those are APTs, but they are not limited to. In the overview of these groups one can

find useful information pertaining to the groups known usage of techniques and what kind of software and capabilities they have. [11]

How to Use MITRE ATT&CK

The use cases of MITRE ATT&CK are varied, but in general it is described as different variations of adversary simulations and penetration testing. A publication by the MITRE Corporation lists a few uses, such as adversary emulation, red teaming, defensive gap assessment and SOC maturity assessment. [12] These have in common that they test and help organizations understand whether they are ready to face a cyber attack from APTs or from methodologies APTs use. This would be useful for any organization believing they could be a target of an APT, and can assist security departments understand where they should place their efforts. Furthermore, the MITRE Corporation lists behavioral analytics development and cyber threat intelligence enrichment, which differ given the fact that they do not test an organization's defensive capabilities. Rather these use cases exist to analyze adversary groups and from that develop more efficient defensive mechanisms. [15]

2.3 SOC

A Cyber Security Operation Center (CSOC) or a Security Operation Center (SOC) which will be used interchangeably throughout this text. According to Schinagle et al. [9] a SOC can be defined as a team of proficient people working with defined processes which is supported by integrated technologies. The SOC mainly focuses on monitoring, incident management, forensic investigation, cyber threats and reporting etc. [9]

For the SOC to be able to monitor ICT systems, applications, infrastructure and services they need their entities such as computers, mobile devices, firewalls, routers, servers, etc. to produce logs. These entities are then referred to as log

sources or assets since they produce logs, which are at first stored locally before being collected and transferred to a centralized repository where the logs from all the different assets are analyzed by the SOC in order to detect incidents. The SOC's main tasks can be divided into the three categories collection, analysis and response which will be elaborated at a later stage.[14]

According to Onwubiko[14] people are a crucial part of the SOC, since technology alone is not enough for the SOC to perform its tasks of monitoring to an adequate level. One of the downsides of people over fully automation is the problem of finding the right people for the job, which also needs to be invested in with training courses etc. which on the other hand could be compared with the continuously updating and development of the automation to keep up with the threats. Another downside of people versus machines is that they get fatigued and therefore are unable to perform to their fullest to any given time. A way to partially mitigate this is to work in shifts, even though this cannot be compared to automations which are always able to perform at the same pace.[14]

A way to optimize the usage of analysts is to use a 3-tier model, based on the analyst's experience and strictness/thoroughness of the expected investigation effort of the event. Where tier 1 analysts are tasked with monitoring and triaging security alerts. The tier 1 analysts are typically the less experienced than the other tiers and mostly rely on predefined playbooks and basic investigation to determine the severity of the alert. While tier 2 analysts are tasked with analyzing and investigating incidents that have been relayed from tier the tier 1 analysts. The tier 2 analysts are typically more experienced than the tier 1 analysts and are therefore able to perform a more complex and in-depth investigation to determine the impact of an incident. Tier 3 analysts is the highest tier of the analysts in the model, consisting of the most experienced and skilled analysts, with a deep understanding of the security tools, threat landscape and the organization's infrastructure. These tier analysts are tasked with handling the most critical and complex incidents that have been relayed by the tier 2 analysts. The tier 3 analysts are able to perform the even more advanced foren-

sic analysis, reverse engineering and threat hunting to identify and remediate security threats. [6]

Collection

Collection of logs is a crucial part of security monitoring. It could be difficult to detect signs of intrusion or if an asset has been compromised without the ability to collect event logs. For a system to be able to collect logs from various assets they first have to be configured to produce logs. These logs should then be sent to a local repository for that system or sub system and thereafter collected in a central repository for all the different systems and sub systems for the SOC to analyze. When configuring the different assets to produce logs it is crucial to ensure that the logs timestamps the logs with a centralized clock source. A centralized clock source gives the ability to detect if the same types of alerts from different assets are related or not.[14]

Analysis

Onwubiko [14] defines analysis as *"The 'brain' behind the CSOC. it is there the logs collected from various assets in the organization are analyzed"*. There are multiple ways to achieve analysis such as manual, fully-automated, or hybrid methods. Manual analysis is when the analysis is done by the security analysts without the help of technologies such as automations. Manual analysis is often done with improvised tools to get the job done, such as for example MS Excel to analyse logs, which could easily become a complex and time consuming task when the log volume becomes too large. Hybrid analysis on the other hand uses the combination of fully automated analysis with the help of human interaction from the security analyst for decision making. Making the hybrid analysis very suitable for protection and continuous monitoring of the organization. While automated analysis utilizes technology for log analysis by performing a series of comprehensive automated tasks without the need for human intervention. This could for example be technology such as Security Information and Event Management(SIEM) systems for analyzing, correlating and/or normalizing data, to be able to swiftly and accurately detect incidents and suspicious activities. In combination with the SIEM one can utilize other technologies such as anomaly

detection systems or web fraud detection to complement the SIEM.[14]

Response

According to Onwubiko [14] there is no longer a question about if an organization will be attacked or have a security breach, but rather a question about when it will happen. One of the SOC's main tasks is incident response. By utilizing a combination of technologies, processes, policies and people to contain, control and mitigate incidents as fast as possible. With the goal of minimizing the impact on the organization, and enable the organization to continue as "normal" while the incident is being managed, instead of having to shut down operation in the meantime. In order for this to work, plans and procedures have to already exist, be well known and regularly tested and updated in advance of an incident.[14]

2.4 SIEM

Security Information and Event Management, or rather SIEM, is a relatively new system designed to offer a platform to make real time monitoring and event analysis possible, as well as logging security data of a time period for both legal and practical reasons. Originally, what we now consider SIEM was two different frameworks, namely Security Event Management and Security Information Management, SEM and SIM respectively.[6]

In recent years, SIEMs have become more or less the backbone of modern SOCs around the world. A SIEM is capable of taking information from multiple source types and aggregate information derived from them into a rich overview whenever an alert is raised. Offering SOC analyst a much bigger picture when performing analysis, reducing time and expertise required to extract and correlate data manually. In recent years, automation technology and AI capabilities have greatly impacted SIEM solutions, making them capable of operating with huge volumes of log data, and to more efficiently recognize threats and false positives.[7]

How Does SIEM Work?

A SIEM aims to gather logs from a wide array of source types within an organization's network and IT systems, this offers an entirely centralized logging solution which again offers several practical possibilities, especially within cybersecurity. Some examples of which source types a SIEM solution can gather data from, are user logs, system antivirus, application logs, network logs, operating systems and even hardware. [3]

From a security perspective, one of the selling points of SIEMs is the fact that it takes information from all these sources and aggregates it into a comprehensive overview for analysts to review whenever an alert is raised. By using correlation rules when monitoring the data flow coming from the different source types, the SIEM can quickly identify irregularities or security incidents. Further it can use pattern recognition and AI to gather related data to create an overview for analysts to investigate, greatly reducing detection and response time for security incidents. [7]

AI and Automations In SIEM

The future of SIEM solution must aim to greatly reduce the amount of false positive alarms created. For example, studies have shown that as little as 29% of alarms are inspected by an analyst, and of those an average of 40% are classified to be false positives. [6] With numbers like this, one may be safe to assume that the likelihood of true positives being missed is relatively high, and as such one must work on technologies to prevent this issue. For example by creating a novel configuration for a SIEM system such as Bryant and Saeidian [6] did by integrating the cyber kill chain. Vastly improving in most metrics compared to a baseline SIEM configuration.

Artificial Inteligence will also play a huge role in managing a SIEM environment in the coming years. With AI becoming much more sophisticated in recent years, it can be applied in alarm analysis, but also speed up correlation and aggregation of data tremendously. With the increase of data from IoT, mobile

and cloud this will be a vital element in the evolution of SIEM systems. [7]

2.5 SOAR

Security Orchestration, Automation and Response or SOAR for short, is a single platform that eneables coordination, execution and automation between tools and users. This single platform technology improves an organization's overall security, by providing the ability to both quickly respond and prevent threats and incidents.[13]

According to Palo Alto Networks [13] SOAR has primarily three main features. Where the first one, threat and vulnerability management (orchestration) is technologies which help manage threats. Secondly SOAR has security operation automation (automation) which covers technology which complement automations and orchestrations in operations. While the third feature is security incident response.[13]

Orchestration and Automation

Security orchestration is the part of SOAR which makes security actions such as incident investigation, resolution and response coordinate within a single infrastructure. Which enables security and non-security tools to work together for both manually and automated tasks, to increase efficiency for processes and security staff. While security automation is the part of SOAR that utilizes machine based executions to investigate, remediate and detect threats without human interaction. Which helps reduce the amount of cases for the SOC team to investigate. Automation could be used to for example detect threats, contain and resolve an incident, or determine whether or not to take action on an incident. [13]

Comparing SOAR and SIEM

SOAR and SIEM are often looked at as similar technologies due to that they both detect security issues and the ability to collect data regarding the security issue. Another similarity between SOAR and SIEM is the ability to notify security personnel about security concerns needed to be addressed.[13]

A difference between SOAR and SIEM is the way data collection and the alerts are collected and generated. While SIEM only sends the alerts to the security analyst, does the SOAR implement automation and response into the path of investigation with the help of automated workflow and playbooks. SOAR also utilizes artificial intelligence to learn and recognize pattern behavior, giving the SOAR the ability to detect or predict threats earlier or even before they happens, based on these patterns. Another difference between SOAR and SIEM is that the SOAR is able to collect alerts from sources that the SIEM is unable to, such as for example alerts from cloud security and IoT devices. [13]

Chapter 3

Methodology

The elected research methodology for this paper is the Design Science Research (DSR) methodology, this chapter includes a summary of what DSR is by exploring Peffers et al's [8] six step plan in conducting a DSR based research project. It also includes why it was chosen for this paper, given how well it matched the proposed task. And how DSR was utilized during the research phase of the project, including the design of artefacts and sub artefacts, choosing the techniques to base research upon, evolving the initial framework to a finished product, and testing the artefacts.

3.1 Design Science Research

Design Science Research(DSR) methodology is a problem solving paradigm which originates/roots from engineering and the science of the artificial. This methodology involves creating and evaluating artifacts such as frameworks, models and software systems etc. to develop knowledge and solve practical problems. [1] According to Peffers et al [8] DSR process can be designed in a 6-step plan which is shown in figure 3.1 below and further explained further down in this chapter.

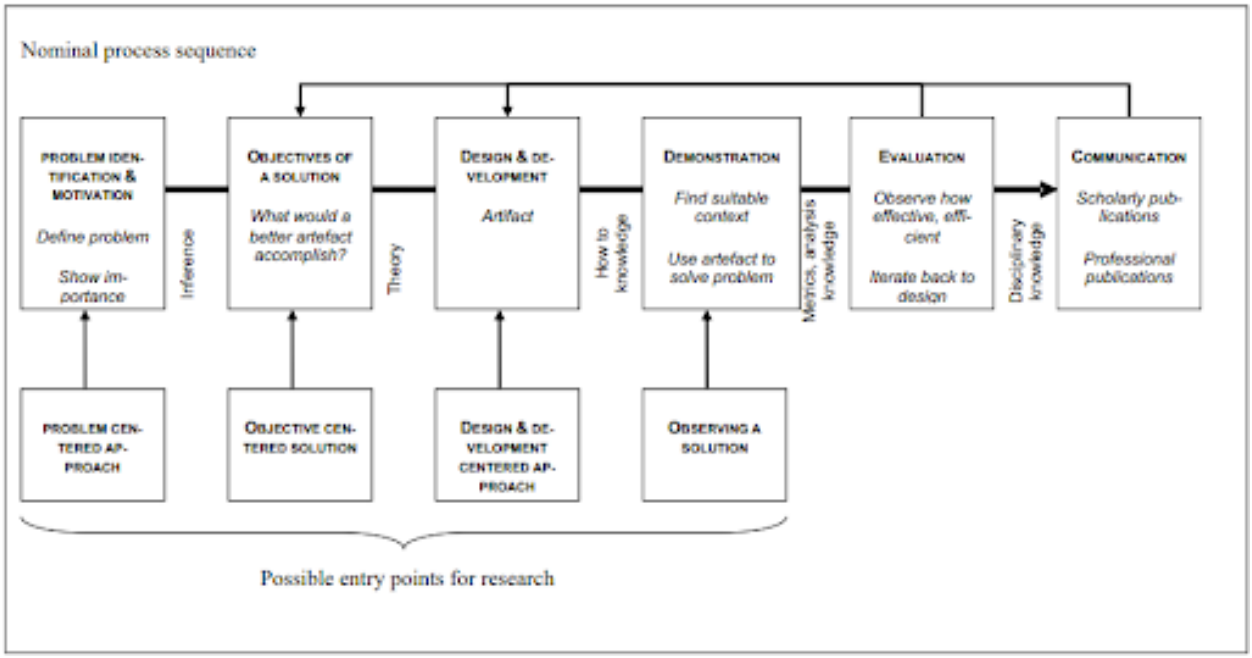

Figure 3.1: DSR process [8]

1. Problem Identification and Motivation

This step involves identifiying the research problem which needs to be solved and establishing a motivation for why the problem is important to solve. As an artifact will be created from the problem definition in the solution, it might be helpful to isolate the problem so that the complexity of the problem can be captured by the solution. [8]

2. Objective of a Solution

Once the problem definition has been identified, the next step is to define the objectives of the solution. These objectives should be aligned with the research problem and a criteria for success should be defined. [8]

3. Design and Development

In this step is where the creation of the artifactual solution happens. This step is about determining the functionality and architectural design of the desired artifact, which satisfies the requirements/boundaries set in the previous steps, before creating the actual artifacts. [8]

4. Demonstration

This step is about demonstrating the efficiency of the artifact for solving the problem. This demonstration could be done by utilizing methods such as case study, simulations or experimentation etc., to gain knowledge and an understanding of how the artifact could be used to solve the problem. [8]

5. Evalutation

This is where the artifact is observed and measured for how effective it supports the solution of the problem. A way of doing this is by comparing the objectives in the solution to the result in step 4 when demonstrating the artifact. The type of comparisons in this step depends on the artifact and the nature of the problem. The evaluation could consist of for example comparing the functionality of the artifact with the solution in step 2.[8]

At the end of this step we could either continue to step 6 "Communication", or we could circle back to step 3 to redesign the artifact to try to improve it. [8]

6. Communication

The final step involves communicating the research such as the problem, importance, artifact and utility etc. In essence, this is the stage where publishing academic article, presenting at conferences. will take place. [8]

Rigor and Relevance

Rigor and relevance are two key principles in DSR which are considered equally important. Rigor is used to ensure that the research findings are reliable and trustworthy. This includes using established research methods and techniques to ensure that the research is conducted in a consistent and systematic manner while addressing any potential biases or errors in the research. While relevance on the other hand is used to ensure that the outcome of the research is useful and practical. Relevance emphasizes on the importance of developing artifacts or solutions which addresses practical problems in a meaningful way. [2]

3.2 DSR Offers a Reasonably Open Road Towards Completion of a Project

Whereas most research methodologies revolve around a single way towards a solution, such as theoretical, experimental or case study, DSR is designed to let the researchers pick their own path, either using any singular method or multiple, in order to reach a goal. [8] It does so by introducing the aforementioned six-step process that spans from creating the idea of an "artefact" to the communication of the "artefact" where the artefact is a method, model, or solution that offers scientific progress toward a problem. This premise was key in our selection of DSR, seeing as when we planned our project we didn't believe there would be any single way to solve it, rather we would use both theoretic and experimental methodologies. Having this level of openness would also be quite beneficial, seeing as this is a reasonably new subject, and could not foresee which kind of problems would arise during the project. This could quite possible force us to change or introduce new methodologies in the process, which would still comply with DSR but not with most other research methodologies.

How DSR Fits the Research

An efficient way of showing how DSR fits well for us, is to discuss some of the steps that define DSR. However, DSR contains a few steps that are quite similar in any methodology, thus we elect to exlude identification & motivations, objective and communication.

Starting with design & development, we start with a theoretical foundation as a solution. The theoretical foundation should consist of information retrieved from the Systematic Literature Review, and other sources such as MITRE themselves and other sources of material that may be relevant to our solution. We realize that because little research is done on this subject, expert opinions from sources other than research reports will be required.

However we know, and it will be a recurring point in this thesis that the theory around this topic is rather slim, as such it will be necessary to enhance this solution through the use of practical experiments. As such the demonstration step in DSR is reached. In this step we will test the model we create by using real data in order to measure its accuracy and efficiency.

Testing the solution will lead to the evaluation step of the DSR. At this stage the solution will be evaluated in order to identify whether or not it is satisfactory in regards to the original goal. Which it may not be, and in such a case, identification of what may be improved and what causes problems will be key, and can be used as we reiterate the process to improve the model.

This introduces the iterative process to our project, we suspect we may have to reiterate several times in order to figure out what will improve our solution and what won't. As visible in figure 3.1, reiteration is a key point in DSR and is to be expected and encouraged.

3.3 Our Implementation of DSR

Coming up with Artefacts and Sub-Artefacts

Seeing as DSR relies on solving a problem through the use of artefacts, it is important to design artefacts that coincide with the research question and the project specification on whomever may use the artefact after its research stage. Considering the vague state of what an artefact really is, it could also be difficult to distinguish which parts of the project may be an artefact and which parts may not. Through meetings with the company we established a foundational understanding on what they wanted our project to accomplish, and the methodology used to reach the results would largely be up to us. Using information gained from these meetings, we were able to envision an idea of what the artefact would accomplish once it was done. Which would be a framework that could be used to create a document explaining the log content needed,

based on any MITRE ATT&CK technique, for an analyst to analyze an alarm without manually extracting information. While also being light enough to not take up an incredible amount of computer power and not cause fatigue.

Knowing what a finished artefact would entail, the next step in the creation of the artefact would be to prepare a methodology for reaching the preferred result with the time frame and resources we had at hand. As a base step, we'd create a basic model to follow. This model would stem from a theoretical foundation, using mostly MITRE as inspiration. Following the MITRE ATT&CK framework we'd use the detection section of each technique to create a foundation of log data to collect for each technique we experimented with. This would then become the most basic version of the artefact possible, from which would evolve with the use of testing on real alarms, input from experts at the company, and further research.

It would not be possible to test this framework without using it on techniques in order to see it's efficiency and challenges, this would generate some sub artefacts, which would essentially be models for a few techniques that could already be automated. These sub artefacts would work for us as a guiding pin in seeing if our research produced results in a favorable way, as well as being results that could be provided to the company immediately after research ended for them to benefit from. Testing and improving these sub artefacts would in turn improve our main artefact by observing what steps would improve the sub artefacts functionality and which steps did not make any difference.

How Techniques Were Chosen

In picking the techniques to be tested there is a variety of considerations to take into account. For example, one could pick three techniques from the same tactic in order to keep the process quite similar for all of them, this would possibly make the process of evaluating log sources easier and more efficient, at the risk of making a framework that may only be efficient for one tactic. Another example could be picking the techniques at random, this way ensuring

no bias would happen and thus ensuring a level of fairness in the experiment. This method however, would run the risk of leaving us with techniques so rarely used, that it would be difficult to source testing material, and it would leave the company with little value. This left us with a few constraints to take into consideration when picking techniques for testing:

- Must not be biased, that is, must contain a level of randomness ensuring we do not pick those that seem easiest or most interesting.
- Must be of relevance to the company, meaning it shouldn't be a technique that contains few to none alarms.
- Must be reasonably common, ensuring enough source material to test, and also ensuring that an analyst that helps us will be familiar with it.
- Should be a part of different tactics, so as to lessen the risk of making a framework only efficient for one tactic.

Identifying a fair selection method seemed to be a difficult task, as the company had not mapped all the alarms to their respective technique in the production version of their SOAR implementation. As luck would have it however, this is a task we had been working on previously. Therefore we had access to a large spreadsheet of almost a thousand alarms, mapped to their respective technique, and were able to sort them by occurrence of technique as depicted in figure 3.2.

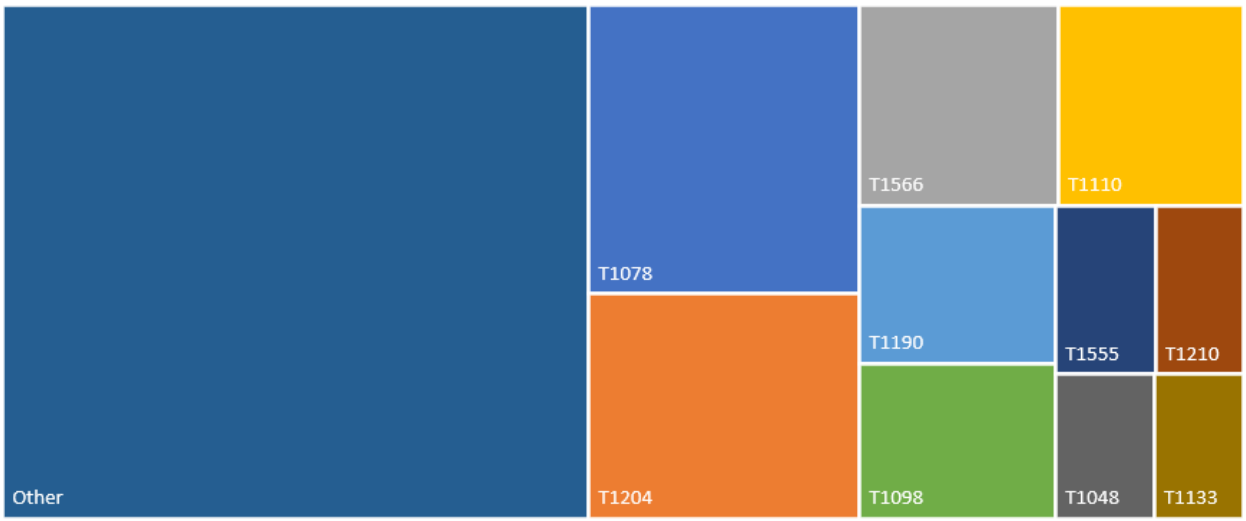

Figure 3.2: Sorted MITRE Techniques

Based on the results generated by the sorted techniques, we were able to select three techniques based on their occurrence:

- Valid Accounts (T1078)
- User Execution (T1204)
- Phishing (T1566)

These three techniques seems to fulfill the requirements that had been set. It's not biased, because it is based on real world material. They are of relevance to the company, this would be confirmed through discussions with an analyst and it is also quite obvious given the fact that they appear so often. Also as previously stated they appear reasonably often, also in the production version of SOAR, so sourcing material is feasible. And they all occur in different tactics namely Initial Access, Execution and Privilege Escalation. Although both Phishing and Valid Accounts occur in the Initial Access tactic as well, we deemed it acceptable given the fact that they also occur in different tactics and all the other factors that weighed positively towards this choice.

3.3.1 Evolving the Framework From a Draft to a Finished Product

In the early stages of the research phase, the framework was little more than gathering intelligence using MITRE ATT&CK to populate data, and testing said data on real alarms in a SOAR environment. The assumption was that MITRE would offer a great bit of the data needed to be tested, and the testing phase would uncover the rest. Although it was well understood that the framework would have to undergo improvements through the research phase, especially after some requirements of the results were set. However, it was unclear in the beginning what those improvements would entail.

Why this approach would not be a good solution can in hindsight be attributed to several factors. Firstly, through discussion and setting requirements for the result, the conclusion became that MITRE's detection sources would require too much data of which a lot would most likely never be used. This issue brought forward the necessity of an additional information gathering process and a discussion process with the goal of narrowing down information sources

needed to a much more specific set of log elements rather than whole log sets. Secondly, the initial idea was that after creating a technique specific model of required data, that would be the final result. However, as attack methodologies constantly keep changing, becomes harder to notice, and more obscured, so would the technique specific model. Thus an extra process of continuously improving the model over time after it was deployed to production as a method to keep it relevant while maintaining a reasonable cost in money, resources and time.

While performing a trial run of the framework in its basic state the conclusion was quickly drawn that the collected data from MITRE would have to be further specified before testing it on real data. The reasoning behind this is that evaluating every element in entire log sources such as MITRE suggests, while possible, is quite infeasible given the fact that the objective is to check each and every log element for its necessity in the model, and remove it if it isn't of any use. Mitigating this issue needed to be a solution that will save time and resources, as such there aren't a lot of ways to proceed. A natural method of saving time is to utilize someone else's work, which is why the step of looking for other information sources is introduced. This, however is not entirely foolproof, as some techniques have lacking information, and one may find information that is dated and incorrect, as such the discussion step was deemed necessary to further optimize the model and decreasing the risk of including wrong elements while still including the correct ones.

Another problem was discovered in hindsight, the issue of what happens to the model after it's created. How it is maintained and updated was previously not thought of but would be of importance. What needed to be solved then, was how to keep the model relevant, using an acceptable level of effort and resources. A naive approach could in this case be to reiterate the process on for example a yearly basis, however, performing such an action on all MITRE technologies would be infeasible as it would require a large amount of time and resources. This approach would also raise the question of how often such a reiteration should take place, more often would leave weaknesses exposed for

a shorter amount of time but would also be very taxing on an organization, and vice versa. Moving away from that idea, a concept well used within SOC environments are "tuning" when necessary. That is, when an issue is discovered within the model, an analyst will either update the model as needed, or relay a request for updating to the relevant personnel. This process however, can quickly cause an overload of information within the model, as elements can be added as necessary while unused elements will likely remain as removal might not be viewed as equally important.

The results can then be identified as a framework including a collection of the elements deemed necessary through testing and requirements. Furthermore it also contains a few elements that in hindsight was deemed constructive and would lead to the models accuracy increasing or relevance, based on experts opinions or theoretical foundation. By performing testing such as discussed in chapter 3.3.3, we were able to evaluate whether or not the resulting model would be satisfactory, and furthermore whether or not the method of reaching the resulting model was satisfactory. Based on findings in the evaluation process we could then discuss adding elements in the process which would make either the model or the process of constructing the model more efficient. Likewise we were also able to exclude certain elements in the event of them making less improvement than the effort it takes to proceed with them, or if they are counter productive. Resulting in tables for each model, containing their respective required log data, and the results of each tested alarm, as depicted in Appendix C. Utilizing reiteration of this process we were able to build a framework deemed to produce results at an acceptable level and for techniques in different tactics in the MITRE framework.

3.3.2 Testing The Artefacts

In order to see progress in the design of the main artefact, it is necessary to test the models created with them. Only through their improvement we'd be able to see improvement of the artefact. Thus testing the main artefact is performed through testing those models, in testing those models we would take 10 alarms in the company's system, belonging to one technique. And put them up against

the model we'd created. Regrettably we were unable to write automations to actually perform the information gathering due to time constraints and a very time consuming learning process that playbooking would be. We had to opt for the second best methodology which was manually investigating the alarm and attempting to reach the same conclusion for the alarm that the analyst had done. Using MITRE, discussion and other theoretical foundation, which is the basic version of the framework that defines the main artefact, we were able to create a model on which log sources we thought would be necessary to solve any alarms belonging to the chosen technique. For each log element we'd discuss whether or not it would be feasible to extract, and whether or not it would give any advantage to the investigation. Likewise if we discovered we were unable to reach the same conclusion as the analyst, based on the sources we had, we'd discuss and research what other sources would be necessary to solve the alarm. Using this process we would add and remove sources as necessary, and reiterate the process to obtain as accurate of a result we could on every alarm in the test.

3.4 Reflection on DSR

Design Science Research seemed like the right choice for this type of research for several reasons. The research project as anticipated in the early stages, utilizes both a theoretical and experimental method to reach the results. There also existed some uncertainty in the early stages, of exactly how the results would be reached, and what type of processes and actions we would need to perform to reach a conclusion, DSR did as expected, let us perform research with a large degree of freedom to choose methodologies as needed, rather than following a strict predefined methodology.

Given these upsides of using DSR for this particular research project, we are quite satisfied in how it guided our research in the right direction. As already mentioned we needed quite a flexible methodology as we were not initially certain of how the research phase would turn out. It did however perform well, due to the iterative nature of DSR which matched the needs of how we needed

to test the models used in researching the framework, as well as the iterative aspect of testing the framework as well. Although research methodologies are many, and some may have worked just as well as DSR did, we believe DSR was the right choice for us in this project.

DSR was indeed the right choice for us, but it did not come without any downsides. One of the major downsides, is the vague nature of DSR, given the oppenness of the methodology, it is not at any given time clear what should be the next step in the research process. It's important to note however, that this degree of openness is also a necessity and was one of the reasons why we picked DSR as previously discussed. Another downside which is closely related to this issue, is the difficulty in interpreting what an artifact is within DSR. There are researchers defining it in different ways, and one is left to somewhat come up with a definition that fits the research.

Chapter 4

Results - A Step by Step Framework

The main artefact of this research project is presented as a step by step framework, which offers an explanation of our proposed methodology of identifying and collecting the data needed to analyze an alarm in a SOC environment based on its MITRE technique. It is an eight step process intended to be aid researchers from the point of deciding a technique, to it's deployment and beyond. Steps in this framework are designed to work as concepts rather than exact instructions, offering the researchers a high degree of freedom in execution, and making the framework available for SOC's utilizing different strategies and platforms. A simple representation of the framework is depicted in figure 4.1, before a detailed version is explained in chapter 4.1

Contents of this framework are the mentioned eight steps, each with elements describing why the step is important, how the step should be executed, which pitfalls to look out for where that is deemed necessary, and images further explaining the step where that is deemed necessary. This way, information is segregated, easily accessible and distinguishable, making it easy for a researcher to quickly identify the necessary information.

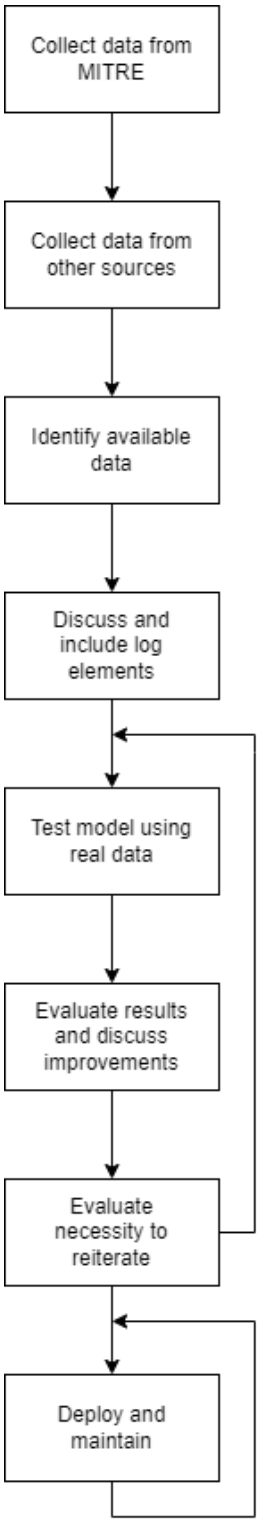

Figure 4.1: Step by Step Process

4.1 Introducing the Framework

This framework is a proof of concept of a methodology to create automated information gathering for alarms, based on their MITRE techniques.

Benefits of framework:

- Cut analysts time spent on retrieving log data and other information required to analyze an alarm.
- Decrease time spent on writing playbooks for specific alarms.
- Should result in a higher degree of order in the system, thus reducing the errors possibly made from confusion and mistakes.
- Basing the system on a well known framework such as MITRE ATT&CK, makes it easier for others than the those creating the automations to take over and understand the system for further development and maintenance.

Requirements of framework:

- Needs to be versatile, that is, work across the whole MITRE ATT&CK framework.
- Must not cause automations to use more computer power than accessible.
- Must rely on technology and log sources available to the company.

4.2 Process of The Framework

In this section we will elaborate each step of the process presented in figure 4.1. For each step we will go through a brief explanation of the step before explaining why the step is necessary, how the step should be performed and where relevant, which pitfalls one should look out for when following this process.

4.2.1 Collect Data From MITRE

Upon choosing a technique to perform this process on, it is a good idea to start by reviewing what information sources MITRE themselves considers useful in order to determine whether or not the alarm is a true positive.

Why?

The MITRE ATT&CK framework is world leading within threat intelligence, and it is written using the collective opinion of cybersecurity experts all over the world. Likewise, there does not seem to be any other information source on this topic as extensive and thorough as the MITRE ATT&CK framework is, adding on to that, the framework is constantly updated and maintained to keep up with the ever evolving threat level in the world's cyber landscape. Utilizing this wealth of information will aid in the process of not only identifying which log sources to include, but also to learn about the technique in order to have a better foundation to discuss elements to include at a more accurate level than what MITRE ATT&CK provides.

How?

The MITRE ATT&CK framework is to a large degree self explanatory, but there are some things one should look at. Some examples include the techniques themselves, how it is utilized, and what MITRE recommends in regards to detection measures.

Pitfalls

MITRE's description is however reasonably vague and in many cases does not specify single elements of data but rather collections of data. If this is all displayed to an analyst, it will leave the risk of an analyst spending large amounts of time searching through all the information, and could lead to fatigue. Still, this doesn't mean that MITRE should be discontinued in the evaluation, because it certainly narrows down what information is useful while offering a wealth of information in discussing why they have been chosen and how attacks in the technique are utilized.

4.2.2 Collect Data From Other Sources

MITRE offers some good information on the general log entries one should look at, however they are somewhat vague and large in size, thus further investigation is necessary. The next step of the process should be to research how techniques or alarms connected to a technique are being analyzed through various online sources.

Why?

Before deciding to build a model of log elements to test with, it can be a good idea to look at sources describing the analyzing process of alarms in regards to the current technique, that is especially true if this process is not performed by, or with an experienced SOC analyst. By performing this step, one may discover actual examples of alarms in the correct technique and how to analyze them. Which will be of immense help in selecting which log data to include in the model, without choosing whole datasets such as MITRE does. This step should not take as much time as testing the models, and thus it should be considered a worthy time investment to further improve the models accuracy by performing this step before performing accuracy tests.

How?

This step of the process will have varying results based on several factors. First of all, some techniques will be utilized more by threat actors than others, naturally leading to said techniques being discussed more in online sources. Likewise, recommended sources such as Microsoft Learn will present their findings based on alarm names and their own category system rather than the technique, which could lead to parts of the technique being well covered while other parts may not.

While Microsoft Learn is inherently difficult to navigate, it contains quite a bit of information in regards to determining the validity of an alarm. This information can in turn be used to identify which log elements would be needed to perform an analysis of the validity of an alert. An effective way to utilize Microsoft

Learn to this extent is to take the alarm name from the SOAR implementation and simply search for it, in quite some cases Microsoft Learn will have a page containing the analysis of said alarm. As an example, some alerts within the Valid Accounts technique could be found here: https://learn.microsoft.com/en-us/defender-cloud-apps/investigate-anomaly-alerts.

Microsoft Learn will however, not contain information on all alarms and techniques, and consequently it is suggested to search for other sources as well. This may bring a lot of relevant information, but it may also not lead to much at all. It is advised to exercise a healthy degree of caution in this process, and make sure sources are reliable.

Pitfalls

As previously mentioned this step will offer results of a varying degree, it is important to perform it with that in mind so as to not hastily choose some sources that have inaccuracies. It may be better to skip this step in the event that no good information can be sourced online, as it may be counterproductive to include the wrong analysis intelligence as opposed to including less intelligence in the model as skipping this step will lead to.

4.2.3 Identify Available Data

Now is a good stage to revisit the information MITRE suggested and identify every available log source and their content, based on MITRE and perhaps other sources identified in step two.

Why?

Upon doing this one will gain an understanding if the log sources available will be enough to continue this process, and will obtain a collection of log sources that will be used in further discussions in the next steps.

How?

Such a process will vary based on systems and information sources that are available. It is however suggested to include an analyst who has a solid grasp on this subject, so as to include every possible information and also to identify which sources may not be feasible to include in an automation and should thus be omitted.

4.2.4 Discuss and Include Log Elements

Nobody knows a company's system and alerts in a SOC environment better than those who analyze them on a daily basis. This stage is a discussion phase where one should make assumptions on which log sources to include, based on information gathered in the previous steps.

Why?

This step is where a techniques model starts to take shape. The result of this step is the foundation of the model, that in later stages will be further tested and evolved. In order to make the next steps as smooth as possible, it is good practice to spend some time creating a proper model now rather than risking having to go through several more iterations of trial and error at a later stage.

How?

In this discussion it can be nice to involve an analyst, and gather as many useful log elements as possible, likewise it is dangerously simple to end up using way more log elements than necessary. One should also note that it is painfully difficult to create a complete model in this stage, and with that in mind one should be weary to not spend too much time trying to perfect it. Thus the results of this stage should be a model of the log sources agreed upon to be necessary to perform an analysis of the alarms in the given technique, which can then be tested in the next stages. Such a model could look like the example in figure 4.2, which is created to serve as a proof of concept and in no way a finished result.

Extracted Data	Alarm 1	Alarm 2	Alarm 3	Alarm 4	Alarm 5	Alarm 6	Alarm 7	Alarm 8	Alarm 9	Alarm 10
Sign in time										
Sign in location										
Sign in user agent										
Success?										
Sign in device										
Failed authentication										
MFA										
VPN?										
Abuseip?										
Calendar?										
Sign in anomalies										

Figure 4.2: Example of table with extracted data

Pitfalls

One of the things that could go wrong at this stage would be that way too many log elements are selected. Should that happen it would definitely provide a great result in terms of the models accuracy, however it would come at the cost of a level of infeasibility both when the model reaches production and an analyst will use it, but also it will cause the testing phase to take an incredibly long time. In addition, if in these discussions there exists doubts and the parts cannot come to an agreement on one or a few log elements, it could prove fruitful to just include them so as to not spend too much time in this phase, those elements usefulness will be identified in the next part either way.

4.2.5 Test Model Using Real Data

Step five in the process is to use the previously built model in order to test it, and there are several ways of performing this task. The main idea however, is to make sure it is tested using real data and it makes sense that the more test material one can use the better the results will become.

Why?

The cyber landscape is ever changing, and the attack methodologies that the threat actors use also changes and becomes more technical and harder to identify. Because of this, it is near impossible to obtain an accurate and lightweight model through theory and reasoning alone. Thus testing the model on real

data will almost definitely uncover further improvements that can improve the model through addition of more log elements or removal of redundant or unused log elements.

How?

While performing these tests, it is important to keep in mind that it is good practice to make a note of which level of significance each log element made in the evaluation of the alarm, this can be used to further improve the model in the next steps. One approach could be to manually use the model and see if it is possible to reach the same conclusion as an analyst who already analyzed the alarm. Using this approach is tedious in the event that a huge sample size is picked. A second approach is largely the same, however, instead of extracting the information the model specifies manually, one would construct a playbook in accordance to the model and use the resulting data to much faster reach a conclusion.

Further on, one can classify the results in 3 classes, if it's difficult or impossible to determine validity of alarm with given information, the result should be classified as failure. Should one be able to tell that it's a true or false positive, but unable to retrieve further information required to provide a decent response, it should be classified as partially acceptable. And at last if one is able to identify that the alarm is a true or false positive and also able to retrieve necessary information for the response it should be classified as acceptable. These results could then be divided into a point system, and calculating an average could determine the accuracy of the model in its current state.

Pitfalls

A point to keep in mind during this step is that one should try to not use alarms with the same name and/or from the same customer, if this is done, there is always the risk of the test only working for a few alarms and not necessarily for the technique itself. Another point is to make sure the sample size is of a large

enough size to be of significance, in order to draw a conclusion of the accuracy and efficiency of the model.

4.2.6 Evaluate Results and Discuss Improvements

It is highly probable that this initial model will have room for improvement, assuming a big enough sample size, that is. At this stage it is time to make evaluations based on previously obtained results.

Why?

Depending on the results obtained in the previous part, there is most likely room for improvement in the model, as previously mentioned. Evaluating these results will result in the possibility of adding new log elements and/or removing unused ones. Which is needed in order to achieve greater accuracy and keep the model as lightweight as possible.

How?

The focus should be on what would be needed to obtain a higher average score, but almost equally important is discussing what can be removed without significantly worsening the results. Adjustments to make here could be removing those log elements that in the previous step proved to never be used in evaluating the alarm. Subsequently it is advised to also identify what caused the failing evaluations to fail, and discuss if any log sources could be included in order to pass them. However, one must remember that including 100 new log elements in the hopes of passing one more alarm in a sample size of 100 alarms may be counter productive given the amount of clutter and unnecessary information the analyst needs to read through in every alarm.

Pitfalls

One of the greatest challenges of this part would lay in deciding the balance between the amount of information the model contains, versus the level of accuracy it provides. This could possibly be solved in discussions with the

analysts who will be using the system every day, because only they know what will lead them to fatigue and confusion, should they have to analyze too much information.

4.2.7 Evaluate Necessity to Reiterate

At this stage in the process, one must consider repeating steps 5 and 6 until a satisfactory accuracy is accomplished.

Why?

As a result of altering the model, unforeseen events may take place. For example, a change believed to solve an issue may not, and a change believed to remove clutter may cause some unforeseen alarms that previously would pass to fail. It is ill advised to believe that a model will function as well or better after changing it, without testing.

How?

Also here some considerations must be taken, it is unlikely for any model to ever reach 100% accuracy. In the event that it does, that may mean the amount of log elements is incredibly high, the sample alarms are too similar or the sample size is not large enough. Furthermore it is quite difficult to know when to stop the reiteration process, and it is largely up to an analyst or any person in charge of performing the construction of the model to choose a satisfactory level of accuracy. It is also possibly the case that it is infeasible to reach a satisfactory level of accuracy, and it is important to move on in the event that several alterations to the model does no longer increase accuracy.

4.2.8 Deploy and Maintain

With a finished model at hand, what remains is to develop it in a playbook and have it released to production. The model however, should never be considered

finished as threat actors and attack vectors constantly change, so must the model.

Why?

There is no telling which new exploits and methods of attack will be used one year from now, and as such there is no telling whether or not the model will be as efficient in the future. By continuously monitoring its performance and making adjustments thereafter one will ensure a future proof model.

How?

There are several methods to go about revising or updating the model. One such method is to have a yearly revision (or any amount of time that seems reasonable, a year could even be too long). Performing this yearly on what will be over a hundred techniques will eventually take up an incredible amount of resources and consequently may be deemed infeasible by many. Therefore it may be more beneficial to do this procedure somewhat differently. Analysts who will rely on this and other models in their work should be able to notice its decrease in effectiveness over time and also where it fails. Having analysts who will report these shortcomings or even fix it themselves will be beneficial in maintaining the model.

Pitfalls

Failure to revise the model will over time decrease its accuracy more and more, as such it may become inefficient to even use it as opposed to another solution. On the other hand, spending too much time and resources could be a costly affair, therefore it is advised to find some middle ground in terms of how much resources to put into it while still keeping it up to date.

Chapter 5

Discussion

This proof of concept framework for creating automatic log gathering based on the MITRE ATT&CK framework is to us a success. With it we were able to make models of how log gathering could be performed based on three different techniques, which performed at a reasonably good level. The framework itself consists of easy to read and segregated bulks of information based on the readers needs, and it offers the reader freedom in methods of execution for most of the steps. However, it is important to note that these findings should be regarded as a proof of concept and further research and testing should be performed before applying this framework in a real word scenario.

5.1 Our Interpretations of Results

For the research to be successful and useful in our opinion, it would be required to fulfill a few points. By reiterating the research question "*How can the MITRE ATT&CK framework be utilized to automate information gathering for security events in a SOC environment?*" we are able to identify a few points that should be fulfilled.

- It should answer the question of if it's possible to create such a framework to fulfill the research question.
- It should utilize a methodology that is feasible to perform for security companies with limited resources.

- The result should in one way or another lessen the load on the analyst by facilitating them to spend more time on analyzing relevant data, and less time on information gathering and analyzing through noise or non-relevant data.

With a set of specified requirements, it is possible to compare our framework with them in order to gauge whether or not the results are useful. For the point of lessening the load on security analysts, without having the framework tested in production with security analysts it is hard to estimate to which degree it increases productivity. However, it is definitely the case that displaying only the relevant fields in log sources that otherwise would be huge, as it will lessen the strain on analyst and increase their efficiency. As for the point of determining if it is actually possible to create such a framework, the results speak for themselves, having tested the framework in small scale we can determine that it is functional, but determining its usefulness and efficiency will require testing on a much larger scale than performed in this research. In terms of feasibility of the framework, it is safe to say that applying it in a real world scenario would require a comprehensive rebuild of most SOC environments, seeing as most SOCs gather alarm information based on other metrics than their MITRE techniques. However once implemented it is extremely feasible to maintain the system, and alterations can easily be performed when it is discovered a lack of log elements.

Furthermore the framework itself which ended up consisting of an 8 step process for a team to follow when creating log data models for any MITRE technique, reached success in the regard that it is relatively short and clearly defined, while still maintaining a level of abstraction that makes it possible to utilize in SOCs of different architecture and design. The frameworks design is intended to segregate different types of information in clearly labeled bulks, making it easy to use whether one would want all the information or just the information needed to perform the task. In a process that is relatively time consuming and repetitive such as this one, one may not need to read all its contents repeatedly

in order to save time, and we believe this segregation strategy aids in relieving teams working with it in that aspect. The choice to keep the framework high level was also done intentionally to reach bigger audiences, we believe that any team working on a project involving this framework will posses the capability to perform the steps even though they are not clearly specified for any architecture or SOC environment. Following this thought process we believe this level of abstraction is beneficial, it also lets those using the framework choose between different methodologies they might have available. One downside of such an approach however, is the possibility that teams may have disagreements in reaching a method of performing this process.

5.2 Acknowledge Limitations

In order to complete the research within a reasonable timeframe, some limitations had to be made which without would cause this project to possibly take years to finish. Given a larger timeframe, testing the log data models could have been done in at a larger scale. By applying the models directly into the SOAR system, it would have facilitated for testing quicker and testing more alarms per technique and more techniques in general. This would however require an extensive knowledge within SOAR and the process of programming its playbooks, as certain parts of this project took longer than expected there was not enough time to perform the testing in this way.

Building on the issue of being unable to perform more extensive tests than executed, we believe that if we did, the results could have become more reliable and accurate. For example, techniques in every relevant tactic were not tested, although it was determined to mainly focus in the three techniques that were tested, more of them would ensure reliability across the tactics which would better satisfy the research question. Likewise, although the tested alarms were chosen at random, a larger sample size would quite possibly uncover more edge cases that could have been rectified, this however only really affects the log data models and not the main results themselves, but it would give a more holistic picture of whether or not the testing process was optimal.

5.3 Share Our Recommendations

For whomever might want to further explore the results produce in this thesis, we'd like to include a couple of recommendations from our perspective. Firstly, it is worth re-mentioning that these results are to be considered a proof of concept and not a complete framework that is ready to be utilized at the production level of a SOC. The reasoning behind this is once again that it has yet to be tested at a major scale and it is to be expected that some flaws will be found during such a process. Furthermore it is not recommended to use the log data models produced during this research, for the same reason as previously stated. Also because different SOCs use different log data sources, and they have a different normalization of said sources, which may produce different results. We therefore recommend to instead follow the same process as shown in figure 4.1 to find the most optimal data and log sources for your SOC.

Further experimenting is recommended however, the framework can certainly become a product with upsides clearly outweighing the downsides. It can greatly aid in the process of keeping information gathering up to date, offering adjustments to easily be applied in real time should the users prefer that. Furthermore it will help maintaining order in codebases that can easily get quite overwhelming and complex in other systems, by dividing the automations into small, easy to read blocks of code with a clear purpose. Rather than having to interpret huge automation processes in order to make a small adjustment. This is especially the case in any SOAR environment where the automation playbooks may become extremely overwhelming. Because of these reasons we believe it is worth expanding this research further in order to thoroughly investigate its capabilities.

Chapter 6

Conclusion

In this research we investigated the usability of the MITRE ATT&CK framework in regards to automating information gathering for security events in a SOC environment. In which the goals were to figure out if it is possible to to identify what data is required to make an automation, whether or not it would be feasible, and how the process of identifying this necessary data could look like, where the focus remained on producing a framework for that process in order to draw a conclusion.

Through theoretical and experimental research we were able to produce a framework detailing a possible process to identify required log data for techniques. Thus answering our research question, concluding that it is indeed possible. Further, this framework provided a detailed set of steps advised to perform in order to make the process as smooth as possible. Although it contains eight steps, it can be summarized into three main parts, being to first establish a theoretical foundation or prototype of an automation, reiterating the process of testing the prototypes and making improvements as needed based on discoveries during the testing, and finally maintaining the finished product by observing its functionality and making adjustments as needed based on analysts using it in real time. Thus this concludes a possible architecture of this framework. Finally in terms of feasibility, one can conclude that it is and quite possibly will be, regardless of methodology, quite time and resource consuming to create these log data models. Although this framework makes an effort to reduce time spent on this process by encouraging the creation of a theoretical foundation before testing and thus reducing the duration of the testing phase,

which can be considered the largest part of the process. Even still, performing this on every technique is quite an undertaking, but maintaining the automations afterwards, following the methods described in the framework, should not add much additional work.

We elected to use the DSR methodology, due to the research we performed may be a little unorthodox, in the sense that it is performed using both theory and experimenting, as such, many research methodologies cannot be used. In addition to DSRs multiple research methodology approach, it also revolves around using artefacts as a means to an end. Interpreting exactly what artefacts are is somewhat challenging, however it can be interpreted as an item being improved or built during the research as a means to reaching the results or the results themselves. By this interpretation, we thought DSR would be even more fitting given the fact that we could identify the framework as a main artefact, and the log data models as sub artefacts.

Our initial expectations for the result differs to some extent, there were certain parts we had assumed would require less work, while some parts we assumed would be more extensive. Initially we did not foresee that the result would end up as a step by step guide, mainly because we had not assumed there would be as many steps to the process as it is. Examples for this are amongst others that we had thought collecting data from MITRE ATT&CK could suffice as a theoretical foundation for the log data models, instead, we discovered that both gathering more data and having a discussion phase to further populate the models before testing was a necessity. Without these steps, one would test this model with very little initial data, and as such one could risk spending quite a lot more time than planned in the testing phase. And the last step which consists of a continuous reiteration of itself where an analyst or another expert updates the models as required in real time, in order to keep the models up to date.

If one looks back at the literature review, one will notice that we were unable to find any research performed on this exact topic, and therefore we can determine that this research as an investigation into a gap within the current existing literature. That's not to say that there aren't any research performed on MITRE ATT&CK, or automation within cybersecurity, or even automation based on MITRE ATT&CK, but there is little to no research exploring the same avenue as this research does. This research, being a proof of concept, cannot necessarily be considered to fill the gap within this topic. However, we hope that this research will be a stepping stone into reaching a higher degree of automated incident handling based on the MITRE ATT&CK framework through future work.

Appendix A

SLR First Search Process Included and Excluded Sources

Title	URL	Search Phrase	Date	Database
99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms	https://www.usenix.org/conference/usenixsecurity22/presentation/alahmadi	“SOC” and “Information Security”	13.02.2023	Google Scholar
A Concept of the Architecture and Creation for SIEM System in Critical Infrastructure	https://link.springer.com/chapter/10.1007/978-3-030-69189-9_13	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
A configurable SoC design for information security	https://ieeexplore.ieee.org/abstract/document/7516998	“SOC” and “Information Security”	15.02.2023	Google Scholar
A Data Triage Retrieval System for Cyber Security Operations Center	https://etda.libraries.psu.edu/catalog/14787txl78	“security operations center”	14.02.2023	Google Scholar
A Deep Neural Network Approach to Tracing Paths in Cybersecurity Investigations	https://ieeexplore.ieee.org/document/9346494	“SOC” and “cybersecurity”	14.02.2023	IEEE Xplore
A Framework Design to Improve and Evaluate the Performance of Security Operation Center (SOC)	https://norma.ncirl.ie/4179/	“SOC” and “IT Security”	13.02.2023	Google Scholar
A Framework for Designing a Security Operations Centre (SOC)	https://ieeexplore.ieee.org/document/7070084	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
A framework for effective threat hunting	https://www.magonlinelibrary.com/doi/abs/10.1016/S1353-4858%2819%2930074-1	“SOC” and “IT Security”	16.02.2023	Google Scholar
A MATURITY CAPABILITY FRAMEWORK FOR SECURITY OPERATION CENTER	https://www.tandfonline.com/doi/full/10.1080/07366981.2023.2159047	“SOC” and “IT Security”	13.02.2023	Google Scholar
A Novel Model for Cybersecurity Economics and Analysis	https://www.webofscience.com/wos/woscc/full-record/WOS:000426119400041	“SOC” and “cybersecurity”	13.02.2023	Web of Science
A Review on the Role of Modern SOC in Cybersecurity Operations	https://ijcsrr.org/wp-content/uploads/2021/05/13-15-2021.pdf	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
A Situational Awareness Dashboard for a Security Operations Center	https://www.proquest.com/openview/4539702c4bc7cf3ba430bfc3d642e6cb/1?pq-origsite=gscholar&cbl=2026366&diss=y	“Att&ck” and “SOAR”	16.02.2023	Google Scholar
A Strategy for Effective Alert Analysis at a Cyber Security Operations Center	https://link.springer.com/chapter/10.1007/978-3-030-04834-1_11	“security operations center”	14.02.2023	Google Scholar
A survey of cybersecurity risk management frameworks	https://www.scopus.com/record/display.uri?eid=2-s2.0-85090098913&origin=resultslist&sort=r-f&src=s&sid=b4bc30bf7627e48d30f12c22e926a55f&sot=b&sdt=b&s=ALL%28%22SOC%22+AND+%22cybersecurity%22%29&sl=40&sessionSearchId=b4bc30bf7627e48d30f12c22e926a55f	“SOC” and “cybersecurity”	17.02.2023	Scopus
A Team-Level Perspective of Human Factors in Cyber Security: Security Operations Centers	https://link.springer.com/chapter/10.1007/978-3-319-60585-2_21	“security operations center”	14.02.2023	Google Scholar
An Automatic Assessment Method of Cyber Threat Intelligence Combined with ATT&CK Matrix	https://www.scopus.com/record/display.uri?eid=2-s2.0-85136637452&origin=resultslist&sort=r-f&src=s&st1=%22Att%26ck%22+and+%22SIEM%22&nlo=&nlr=&nls=&sid=07a4ace08c712ba0d1978bdd36ecac26&sot=b&sdt=cl&cluster=scofreetoread%2c%22all%22%2ct&sl=43&s=ALL%28%22Att%26ck%22+and+%22SIEM%22%29+AND+PUBYEAR+%3e+2012&relpos=2&citeCnt=0&searchTerm=	“Att&ck” and “SIEM”	17.02.2023	Scopus
An Easy-to-use Framework to Build and Operate AI-based Intrusion Detection for In-situ Monitoring	https://ieeexplore.ieee.org/document/9546868	“SOC” and “cybersecurity”	14.02.2023	IEEE Xplore
An integrative review and analysis of cybersecurity research: Current state and future directions	https://publications.aaahq.org/jis/article-abstract/35/1/155/962/An-Integrative-Review-and-Analysis-of	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
An Overview of MITRE Cyber Situational Awareness Solutions	https://apps.dtic.mil/sti/citations/AD1107812	“MITRE” and “SIEM”	14.02.2023	Google Scholar

Title	URL	Search Phrase	Date	Database
Analysis of Cybersecurity Standard and Framework Components	https://d1wqtxts1xzle7.cloudfront.net/78607584/426-libre.pdf?1642083233=&response-content-disposition=inline%3B+filename%3DAnalysis_of_Cybersecurity_Standard_and_F.pdf&Expires=1676465879&Signature=Uf1KWwFIlyOgGVadfJkiX4MIsprp57~nADGwTm1nzXutexDzuN2zWpofYOT1R8uj2PJsIeUzKGJu2yNTBtHnK5QzV32b4bGnV1ANmxh~hvhjM~Wwvci9Jcn6LPGZ1df2Qy-QJn4VVcOtgqYsbfkdOgw~kVaY9cFfZsrrhURKnZhwZnlgInSjpPfMzJ-b6G9uu51KCAd5aM5ZtlaT4GGlgJTr-m~nvA06MnGHTGsNSnoBL7AVJNNqo3S9d5qEDr2lqZCDJLhB0LPt7NcdPQdJaYSZnUc-gnrJjl2yZggIzmYdQJQSOJmZVLduOzJCWXb30~gPseLLDsR5iwaarb-8YA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
Analysis of SIEM Systems and Their Usage in Security Operations and Security Intelligence Centers	https://link.springer.com/chapter/10.1007/978-3-319-63940-6_40	“security operations center”	14.02.2023	Google Scholar
Analysis of the Functionalities of a Shared ICS Security Operations Center	https://ieeexplore.ieee.org/document/9010607	“SOC” and “cybersecurity”	14.02.2023	IEEE Xplore
Analysis on Security Orchestration Automation and Response (SOAR) platforms for Security Operation Centers	https://dione.lib.unipi.gr/xmlui/handle/unipi/14560	“MITRE” and “SOAR”	14.02.2023	Google Scholar
Analytical visualization techniques for security information and event management	https://ieeexplore.ieee.org/abstract/document/6498600	“MITRE” and “SIEM”	14.02.2023	Google Scholar
Anomaly Detection by Recombining Gated Unsupervised Experts	https://www.webofscience.com/wos/woscc/full-record/WOS:000867070907044	“SOC” and “IT Security”	13.02.2023	Web of Science
Application of the Metric Learning for Security Incident Playbook Recommendation	https://ieeexplore.ieee.org/abstract/document/9507632	“MITRE” and “SOAR”	14.02.2023	Google Scholar
ART: Automated Reclassification for Threat Actors based on ATT&CK Matrix Similarity	https://ieeexplore.ieee.org/document/9559514	“MITRE” and “SOC”	14.02.2023	IEEE Xplore
Assessing MITRE ATT&CK Risk Using a Cyber-Security Culture Framework	https://www.webofscience.com/wos/woscc/full-record/WOS:000650792200001	“security operations center”	14.02.2023	Web of Science
ATT&CK Behavior Forecasting based on Collaborative Filtering and Graph Databases	https://ieeexplore.ieee.org/document/10032036	“Att&ck” and “SOC”	14.02.2023	IEEE Xplore
Automated Identification of Cyber Threat Scenarios	https://is.muni.cz/th/s4byh/RIGO_Sadlek.pdf	“Att&ck” and “SOAR”	14.02.2023	Google Scholar
Automatic Narrative Summarization for Visualizing Cyber Security Logs and Incident Reports	https://www.scopus.com/record/display.uri?eid=2-s2.0-85123854170&origin=resultslist&sort=r-f&src=s&st1=%22MITRE%22+and+%22SOC%22&nlo=&nlr=&nls=&sid=c6c7bd03dca1d1baf5352569c5b90cea&sot=b&sdt=cl&cluster=scofreetoread%2c%22all%22%2ct&sl=22&s=ALL%28%22MITRE%22+and+%22SOC%22%29&relpos=55&citeCnt=3&searchTerm=	“MITRE” and “SOC”	17.02.2023	Scopus
Automating Reasoning with ATT (and) CK	https://apps.dtic.mil/sti/citations/AD1088924	“MITRE” and “SIEM”	14.02.2023	Google Scholar
Automation in cyber security	https://www.theseus.fi/handle/10024/503899	“MITRE” and “SOAR”	14.02.2023	Google Scholar
Automation in the Cybersecurity Incident Handling Process	https://webthesis.biblio.polito.it/24503/	“Att&ck” and “SOC”	14.02.2023	Google Scholar
Automation of Risk-Based Vulnerability Management Based on a Cyber Kill Chain Model	https://ceur-ws.org/Vol-3056/paper-14.pdf	“Att&ck” and “triage”	14.02.2023	Google Scholar
Big Fish, Little Fish, Critical Infrastructure: An Analysis of Phineas Fisher and the 'Hacktivist' Threat to Critical Infrastructure	https://www.webofscience.com/wos/woscc/full-record/WOS:000847358100025	“MITRE” and “SOC”	14.02.2023	Web of Science
Classification of Security Operation Centers	https://ieeexplore.ieee.org/abstract/document/6641054	“SOC” and “Information Security”	15.02.2023	Google Scholar
CoCoa: An Ontology for Cybersecurity Operations Centre Analysis Process	https://ieeexplore.ieee.org/document/8551486	“SOC” and “cybersecurity”	14.02.2023	IEEE Xplore

Title	URL	Search Phrase	Date	Database
Cognitive security: A comprehensive study of cognitive science in cybersecurity	https://www.sciencedirect.com/science/article/pii/S2214212618307804	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Common and Best Practices for Security Operations Centers: Results of the 2019 SOC Survey	https://pages.siemplify.co/rs/182-SXA-457/images/Survey_SOC-2019_Siemplify.pdf	“SOC” and “IT Security”	16.02.2023	Google Scholar
Conceptual Model for a Shared Cybersecurity Operations Center for ICS	https://link.springer.com/chapter/10.1007/978-3-030-90321-3_40	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Conscious Machines for Autonomous Agents and Cybersecurity	https://www.webofscience.com/wos/woscc/full-record/WOS:000828673500012	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Continuous improvement on maturity and capability of Security Operation Centres	https://www.webofscience.com/wos/woscc/full-record/WOS:000607087400001	“SOC” and “Information Security”	13.02.2023	Web of Science
Customized access log classifier for cybersecurity incident detection	https://www.webofscience.com/wos/woscc/full-record/WOS:000592093000007	“security operations center”	14.02.2023	Web of Science
Cyber Security Operations Centre Concepts and Implementation	https://www.igi-global.com/chapter/cyber-security-operations-centre-concepts-and-implementation/253664	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy	https://ieeexplore.ieee.org/document/7166125	“SOC” and “Information Security”	14.02.2023	IEEE Xplore
Cyber Threat Modeling: Survey, Assessment, and Representative Framework	https://apps.dtic.mil/sti/citations/AD1108051	“MITRE” and “Triage”	16.02.2023	Google Scholar
CyberOps: Situational Awareness in Cybersecurity Operations	https://arxiv.org/abs/2202.03687	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Cybersecurity Operations Center	https://www.dut.edu.ua/uploads/l_1717_91042607.pdf	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
Demonstration of the Cybersecurity Framework through Real-World Cyber Attack	https://www.webofscience.com/wos/woscc/full-record/WOS:000540654600003	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Demystifying analytical information processing capability: The case of cybersecurity incident response	https://www.sciencedirect.com/science/article/pii/S0167923620302311	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Design and Development of Automated Threat Hunting in Industrial Control Systems	https://www.webofscience.com/wos/woscc/full-record/WOS:000821801200129	“MITRE” and “SOC”	14.02.2023	Web of Science
Early Detection of Cybersecurity Threats Using Collaborative Cognition	https://www.webofscience.com/wos/woscc/full-record/WOS:000519942300042	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Effectiveness of cybersecurity audit	https://www.scopus.com/record/display.uri?eid=2-s2.0-85123639413&origin=resultslist&sort=r-f&src=s&sid=b4bc30bf7627e48d30f12c22e926a55f&sot=b&sdt=b&s=ALL%28%22SOC%22+AND+%22cybersecurity%22%29&sl=40&sessionSearchId=b4bc30bf7627e48d30f12c22e926a55f	“SOC” and “cybersecurity”	17.02.2023	Scopus
Enhancing Collaboration Between Security Analysts in Security Operations Centers	https://www.webofscience.com/wos/woscc/full-record/WOS:000876630700012	“security operations center”	14.02.2023	Web of Science
Enhancing the STIX Representation of MITRE ATT&CK for Group Filtering and Technique Prioritization	https://books.google.no/books?hl=en&lr=&id=VtB2EAAAQBAJ&oi=fnd&pg=PA385&dq=%E2%80%9CMITRE%E2%80%9D+and+%E2%80%9CSOC%E2%80%9D&ots=4QKD46BgmP&sig=rkIjfN27Bau5meQeMHAeSGqqO6M&redir_esc=y#v=onepage&q=%E2%80%9CMITRE%E2%80%9D%20and%20%E2%80%9CSOC%E2%80%9D&f=false	“MITRE” and “SOC”	14.02.2023	Google Scholar
Exploring the applicability of SIEM technology in IT security	https://openrepository.aut.ac.nz/handle/10292/7688	“SOC” and “IT Security”	13.02.2023	Google Scholar
Exploring VirusTotal for security operations alert triage automation	https://www.theseus.fi/handle/10024/704502	“MITRE” and “SOAR”	14.02.2023	Google Scholar
Extracting Cybersecurity Related Linked Data from Text	https://www.webofscience.com/wos/woscc/full-record/WOS:000330582900039	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Extracting Rich Semantic Information about Cybersecurity Events	https://www.webofscience.com/wos/woscc/full-record/WOS:000554828705017	“SOC” and “cybersecurity”	13.02.2023	Web of Science

Title	URL	Search Phrase	Date	Database
Factors Contributing to the Success of Information Security Management Implementation	https://pdfs.semanticscholar.org/7787/a372b03833997f08aa44d1b2e8d6f92c7726.pdf	“SOC” and “Information Security”	15.02.2023	Google Scholar
From logs to Stories: Human-Centred Data Mining for Cyber Threat Intelligence	https://www.webofscience.com/wos/woscc/full-record/WOS:000524755900003	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Grasp on next generation security operation centre (NGSOC): Comparative study	https://ijnaa.semnan.ac.ir/article_5145.html	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
Heated Alert Triage (HeAT): Network-Agnostic Extraction of Cyber Attack Campaigns	https://ceur-ws.org/Vol-3095/paper3.pdf	“MITRE” and “Triage”	14.02.2023	Google Scholar
How to Build a SOC on a Budget	https://ieeexplore.ieee.org/document/9850281	“security operations center”	14.02.2023	IEEE Xplore
Improving information and task coordination in cyber security operation centers	https://www.proquest.com/openview/0f2f7f6f5ae63ebdfaa19bcd259c66a4/1?pq-origsite=gscholar&cbl=51908	“security operations center”	14.02.2023	Google Scholar
Improving SIEM alert metadata aggregation with a novel kill-chain based classification model	https://www.sciencedirect.com/science/article/pii/S016740482030095X	“Att&ck” and “triage”	14.02.2023	Google Scholar
Improving threat detection with a detection development life cycle	https://www.ingentaconnect.com/content/hsp/jcs/2021/00000005/00000002/art00003	“MITRE” and “SOAR”	16.02.2023	Google Scholar
Information alignment and visualization for security operations center teams	https://dl.acm.org/doi/abs/10.5555/2775498.2775517	“security operations center”	14.02.2023	Google Scholar
Information Security Issues Analysis and Solution	https://www.scopus.com/record/display.uri?eid=2-s2.0-85145436160&origin=resultslist&sort=r-f&src=s&sid=b4bc30bf7627e48d30f12c22e926a55f&sot=b&sdt=cl&s=ALL%28%22SOC%22+AND+%22Information+Security%22%29&sl=40&sessionSearchId=b4bc30bf7627e48d30f12c22e926a55f	“SOC” and “Information Security”	17.02.2023	Scopus
Information security knowledge sharing in organizations: Investigating the effect of behavioral information security governance and national culture	https://www.sciencedirect.com/science/article/pii/S0167404814000339	“SOC” and “Information Security”	13.02.2023	Google Scholar
Information visualization metrics and methods for cyber security evaluation	https://ieeexplore.ieee.org/document/6578846	“SOC” and “Information Security”	14.02.2023	IEEE Xplore
Insider Threat Cybersecurity Framework Webtool & Methodology: Defending Against Complex Cyber-Physical Threats	https://www.webofscience.com/wos/woscc/full-record/WOS:000674762400028	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Integrating a security operations centre with an organization's existing procedures, policies and information technology systems	https://ieeexplore.ieee.org/abstract/document/8601251	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Investigating Proactive Digital Forensics Leveraging Adversary Emulation	https://www.webofscience.com/wos/woscc/full-record/WOS:000858034300001	“MITRE” and “Triage”	14.02.2023	Web of Science
Learning From Experts' Experience: Toward Automated Cyber Security Data Triage	https://www.webofscience.com/wos/woscc/full-record/WOS:000459697700057	“security operations center”	14.02.2023	Web of Science
Management and Monitoring Security Events in a Business Organization - SIEM system	https://ieeexplore.ieee.org/abstract/document/9803428	“SOC” and “IT Security”	13.02.2023	Google Scholar
Matched and Mismatched SOCs: A Qualitative Study on Security Operations Center Issues	https://www.webofscience.com/wos/woscc/full-record/WOS:000509760700118	“security operations center”	14.02.2023	Web of Science
MATE: Summarizing Alerts to Interpretable Outcomes with MITRE ATT&CK	https://ieeexplore.ieee.org/abstract/document/10020587	“MITRE” and “SOC”	14.02.2023	Google Scholar
MAVEN information security governance, risk management, and compliance (GRC): Lessons learned	https://ieeexplore.ieee.org/document/6836516	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
Measuring the technical performance of a security operations center	https://jyx.jyu.fi/handle/123456789/84367	“security operations center”	14.02.2023	Google Scholar
MITRE ATT and CK(trademark): Design and Philosophy	https://apps.dtic.mil/sti/citations/AD1108016	“MITRE” and “SOC”	16.02.2023	Google Scholar
Monitoring and Improving Managed Security Services inside a Security Operation Center	https://spectrum.library.concordia.ca/id/eprint/980759/	“SOC” and “IT Security”	13.02.2023	Google Scholar
Near Real-time Learning and Extraction of Attack Models from Intrusion Alerts	https://arxiv.org/abs/2103.13902	“Att&ck” and “SOC”	14.02.2023	Google Scholar

Title	URL	Search Phrase	Date	Database
Next Generation SOC: Automations and Machine Learning in Cybersecurity	https://webthesis.biblio.polito.it/25397/	“Att&ck” and “SOAR”	14.02.2023	Google Scholar
NIST CyberSecurity Framework Compliance A Generic Model for Dynamic Assessment and Predictive Requirements	https://www.webofscience.com/wos/woscc/full-record/WOS:000399004200053	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Offensive Security: Towards Proactive Threat Hunting via Adversary Emulation	https://www.webofscience.com/wos/woscc/full-record/WOS:000696652500001	“MITRE” and “SIEM”	14.02.2023	Web of Science
Operational security, threat intelligence & distributed computing: the WLCG Security Operations Center Working Group	https://www.webofscience.com/wos/woscc/full-record/WOS:000570241300127	“security operations center”	14.02.2023	Web of Science
Optimizing Alert Data Management Processes at a Cyber Security Operations Center	https://link.springer.com/chapter/10.1007/978-3-030-30719-6_9	“security operations center”	14.02.2023	Google Scholar
Pushing the Limits in Event Normalisation to Improve Attack Detection in IDS/SIEM Systems	https://ieeexplore.ieee.org/document/6824575	“MITRE” and “SIEM”	14.02.2023	IEEE Xplore
Quantifying and Analyzing Information Security Risk from Incident Data	https://link.springer.com/chapter/10.1007/978-3-030-36537-0_7	“SOC” and “Information Security”	13.02.2023	Google Scholar
RADAR: Effective Network-based Malware Detection based on the MITRE ATT&CK Framework	https://arxiv.org/abs/2212.03793	“MITRE” and “SOC”	14.02.2023	Google Scholar
RECONSTRUCTING ALERT TREES FOR CYBER TRIAGE	https://profsandhu.com/ics/2022%20Eric%20Ficke.pdf	“MITRE” and “Triage”	16.02.2023	Google Scholar
Regression-Based Attack Chain Analysis and Staffing Optimization for Cyber Threat Detection	https://www.proquest.com/openview/3803b4642c735a04f124e2095a8ed99d/1?cbl=18750&diss=y&pq-origsite=gscholar&parentSessionId=Id1vPwzcF1vYFxTjFzIHP%2Fv97CUGtqq7GBTYbS%2FWD9E%3D	“Att&ck” and “SOC”	16.02.2023	Google Scholar
RelExt: Relation Extraction using Deep Learning approaches for Cybersecurity Knowledge Graph Improvement	https://www.webofscience.com/wos/woscc/full-record/WOS:000555683800150	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Requirement semi-formalization methodology for SoC design	https://ieeexplore.ieee.org/document/7401686	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
S.P.O.O.F Net: Syntactic Patterns for identification of Ominous Online Factors	https://ieeexplore.ieee.org/document/8424657	“SOC” and “Information Security”	14.02.2023	IEEE Xplore
SAIBERSOC: Synthetic Attack Injection to Benchmark and Evaluate the Performance of Security Operation Centers	https://www.webofscience.com/wos/woscc/full-record/WOS:000684258200011	“security operations center”	14.02.2023	Web of Science
SAMIIT: Spiral attack model in IIoT mapping security alerts to attack life cycle phases	https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/ICS2018.2	“Att&ck” and “SOC”	14.02.2023	Google Scholar
Scalable SoC trust verification using integrated theorem proving and model checking	https://ieeexplore.ieee.org/document/7495569	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
Security analytics tools and implementation success factors: Instrument development using Delphi approach and exploratory factor analysis	https://www.proquest.com/openview/92e5e8a617d80f03d8be93c6204a709d/1?cbl=18750&pq-origsite=gscholar&parentSessionId=LmXolnEp3FMN79dZYlb8PZV%2BPS4J14hPUbhmhU252R0%3D	“SOC Efficiency” and “MITRE”	14.02.2023	Google Scholar
Security Concerns Towards Security Operations Centers	https://ieeexplore.ieee.org/document/8440963	“security operations center”	14.02.2023	IEEE Xplore
Security information and event management (SIEM): Analysis, trends, and usage in critical infrastructures	https://www.scopus.com/record/display.uri?eid=2-s2.0-85109535578&origin=resultslist&sort=r-f&src=s&st1=%22SOC%22+and+%22cybersecurity%22&nlo=&nlr=&nls=&sid=8fe803ce9d7ed7541ca5396ba491f7e8&sot=b&sdt=cl&cluster=scofreetoread%2c%22all%22%2ct&sl=49&s=ALL%28%22SOC%22+and+%22cybersecurity%22%29+AND+PUBYEAR+%3e+2012&relpos=143&citeCnt=14&searchTerm=	“SOC” and “cybersecurity”	14.02.2023	Scopus
Security Operations Center (SOC)	https://www.researchgate.net/profile/Manfred-Vielberth-2/publication/349312209_Security_Operations_Center_SOC/links/602a4428299bf1cc26c861fe/Security-Operations-Center-SOC.pdf	“SOC” and “IT Security”	16.02.2023	Google Scholar
Security Operations Center: A Framework for Automated Triage, Containment and Escalation	https://www.scirp.org/journal/paperinformation.aspx?paperid=103116	“SOC” and “Information Security”	13.02.2023	Google Scholar

Title	URL	Search Phrase	Date	Database
Security Operations Center: A Systematic Study and Open Challenges	https://ieeexplore.ieee.org/document/9296846	“security operations center”	14.02.2023	IEEE Xplore
Security Operations Centers for Information Security Incident Management	https://ieeexplore.ieee.org/document/7575854	“SOC” and “Information Security”	14.02.2023	IEEE Xplore
Security operations centre: situation awareness, threat intelligence and cybercrime	https://ieeexplore.ieee.org/document/8057355	“security operations center”	18.02.2023	IEE Xplore
Security usability principles for vulnerability analysis and risk assessment	https://www.webofscience.com/wos/woscc/full-record/WOS:000253605800027	“SOC” and “IT Security”	13.02.2023	Web of Science
SHADEWATCHER: Recommendation-guided Cyber Threat Analysis using System Audit Records	https://ieeexplore.ieee.org/document/9833669	“cybersecurity” and “automated alarm analysis”	14.02.2023	IEEE Xplore
SIEM selection criteria for an efficient contextual security	https://ieeexplore.ieee.org/document/8072035	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
Siem-Enabled Cyber Event Correlation (What And How)	https://apps.dtic.mil/sti/citations/AD1065276	“MITRE” and “SIEM”	16.02.2023	Google Scholar
SmartValidator: A framework for automatic identification and classification of cyber threat data	https://www.scopus.com/record/display.uri?eid=2-s2.0-85127644255&origin=resultslist&sort=r-f&src=s&st1=%22MITRE%22+and+%22SOC%22&nlo=&nlr=&nls=&sid=c6c7bd03dca1d1baf5352569c5b90cea&sot=b&sdt=cl&cluster=scofreetoread%2c%22all%22%2ct&sl=22&s=ALL%28%22MITRE%22+and+%22SOC%22%29&relpos=48&citeCnt=3&searchTerm=	“MITRE” and “SOC”	17.02.2023	Scopus
SOAR Playbook Implementation - Incident Deduplication and Its Effects	https://www.theseus.fi/bitstream/handle/10024/354155/thesis_purujoki_jani.pdf?sequence=2	“MITRE” and “SOAR”	16.02.2023	Google Scholar
SOC as a Service : a user centric approach for Network Security Monitoring	https://www.researchgate.net/profile/Nicolas-Greneche/publication/337902339_SOC_as_a_Service_A_User_Centric_Approach_for_Network_Security_Monitoring/links/5df17755a6fdcc28371a2ff1/SOC-as-a-Service-A-User-Centric-Approach-for-Network-Security-Monitoring.pdf	“SOC” and “IT Security”	16.02.2023	Google Scholar
SOC ATTACKER CENTRIC - Analysis of a prevention oriented SOC	https://www.utupub.fi/bitstream/handle/10024/174290/Ioris_Mirko_Thesis.pdf?sequence=1	“Att&ck” and “SOC”	16.02.2023	Google Scholar
SOC Critical Path: A Defensive Kill Chain Model	https://www.webofscience.com/wos/woscc/full-record/WOS:000753420500001	“security operations center”	14.02.2023	Web of Science
SoC Trust Validation Using Assertion-Based Security Monitors	https://ieeexplore.ieee.org/document/9424363	“SOC” and “IT Security”	14.02.2023	IEEE Xplore
SOC- and SIC-Based Information Security Monitoring	https://link.springer.com/chapter/10.1007/978-3-319-56538-5_37	“SOC” and “Information Security”	13.02.2023	Google Scholar
Success factors for cyber security operation center (SOC) establishment	https://eudl.eu/doi/10.4108/eai.18-7-2019.2287841	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
Tactical provenance analysis for endpoint detection and response systems	https://ieeexplore.ieee.org/abstract/document/9152771	“MITRE” and “Triage”	14.02.2023	Google Scholar
Temporal Understanding of Cybersecurity Threats	https://www.webofscience.com/wos/woscc/full-record/WOS:000631254500019	“SOC” and “cybersecurity”	13.02.2023	Web of Science
The Applicability of a SIEM Solution: Requirements and Evaluation	https://ieeexplore.ieee.org/document/8795405	“SOC” and “cybersecurity”	18.02.2023	IEE Xplore
THE EVOLUTION OF NETWORK BASED CYBERSECURITY NORMS: An Analytical Narrative	https://www.webofscience.com/wos/woscc/full-record/WOS:000380450500088	“SOC” and “cybersecurity”	13.02.2023	Web of Science
THE HUMAN FACTOR CAPABILITIES IN SECURITY OPERATION CENTER (SOC)	https://www.tandfonline.com/doi/full/10.1080/07366981.2021.1977026	“SOC” and “cybersecurity”	15.02.2023	Google Scholar
The Intelligent Process Lifecycle of Active Cyber Defenders	https://dl.acm.org/doi/full/10.1145/3499427	“Att&ck” and “SIEM”	14.02.2023	Google Scholar
The Next Gen Security Operation Center	https://ieeexplore.ieee.org/document/9418136	“SOC” and “Information Security”	14.02.2023	IEEE Xplore
The next generation cognitive security operations center: network flow forensics using cybersecurity intelligence	https://www.mdpi.com/2504-2289/2/4/35	“SOC” and “cybersecurity”	13.02.2023	Google Scholar
The problem with (most) network detection and response	https://www.magonlinelibrary.com/doi/abs/10.1016/S1353-4858%2820%2930104-5	“MITRE” and “SIEM”	16.02.2023	Google Scholar

Title	URL	Search Phrase	Date	Database
The Quantification and Analysis of Cyber-Security Operations Center Vulnerability Data	https://www.proquest.com/openview/813ba2abe35f118343dfbdf9976920ff/1?pq-origsite=gscholar&cbl=18750	“security operations center”	14.02.2023	Google Scholar
The Seven Golden Principles of Effective Anomaly-Based Intrusion Detection	https://www.scopus.com/record/display.uri?eid=2-s2.0-85114710212&origin=resultslist&sort=r-f&src=s&st1=%22Att%26ck%22+and+%22SIEM%22&nlo=&nlr=&nls=&sid=07a4ace08c712ba0d1978bdd36ecac26&sot=b&sdt=cl&cluster=scofreetoread%2c%22all%22%2ct&sl=43&s=ALL%28%22Att%26ck%22+and+%22SIEM%22%29+AND+PUBYEAR+%3e+2012&relpos=1&citeCnt=0&searchTerm=	“Att&ck” and “SIEM”	17.02.2023	Scopus
Threat Hunting as Cyber Security Baseline in the Next-Generation Security Operations Center	https://ieeexplore.ieee.org/document/9653361	“security operations center”	14.02.2023	IEEE Xplore
Threat Management Based on Information About Vulnerabilities	https://is.muni.cz/th/vpeip/diploma_thesis.pdf	“Att&ck” and “triage”	16.02.2023	Google Scholar
Threshold-Based Widespread Event Detection	https://www.webofscience.com/wos/woscc/full-record/WOS:000565234200038	“SOC” and “cybersecurity”	13.02.2023	Web of Science
Time is of the Essence: Machine Learning-based Intrusion Detection in Industrial Time Series Data	https://www.webofscience.com/wos/woscc/full-record/WOS:000465766800001	“SOC” and “IT Security”	13.02.2023	Web of Science
Towards an Automated Dissemination Process of Cyber Threat Intelligence Data using STIX	https://ieeexplore.ieee.org/abstract/document/9631850	“Att&ck” and “SOC”	14.02.2023	Google Scholar
Towards Automated Threat-Informed Cyberspace Defense	https://www.duo.uio.no/handle/10852/88302	“Att&ck” and “SOAR”	14.02.2023	Google Scholar
Towards Automatic Property Generation for SoC Security Verification	https://ieeexplore.ieee.org/document/10031448	“SOC” and “cybersecurity”	14.02.2023	IEEE Xplore
Towards automation of threat modeling based on a semantic model of attack patterns and weaknesses	https://arxiv.org/abs/2112.04231	“Att&ck” and “SIEM”	14.02.2023	Google Scholar
Towards Mitigation of Data Exfiltration Techniques Using the MITRE ATT&CK Framework	https://www.scopus.com/record/display.uri?eid=2-s2.0-85132732777&origin=resultslist&sort=r-f&src=s&sid=b4bc30bf7627e48d30f12c22e926a55f&sot=b&sdt=cl&s=ALL%28%22MITRE%22+AND+%22SOC%22%29&sl=40&sessionSearchId=b4bc30bf7627e48d30f12c22e926a55f	“MITRE” and “SOC”	17.02.2023	Scopus
Traceability for Adaptive Information Security in the Cloud	https://www.webofscience.com/wos/woscc/full-record/WOS:000392940500131	“SOC” and “Information Security”	13.02.2023	Web of Science
Understanding and Enabling Tactical Situational Awareness in a Security Operations Center	https://link.springer.com/chapter/10.1007/978-3-030-52581-1_10	“security operations center”	14.02.2023	Google Scholar
Understanding situation awareness in SOCs, a systematic literature review	https://www.webofscience.com/wos/woscc/full-record/WOS:000914752600001	“security operations center”	14.02.2023	Web of Science
USING INDICATORS OF COMPROMISE TO AUTOMATE INCIDENT TRIAGE. PROOF OF CONCEPT.	https://digikogu.taltech.ee/en/Download/835d07ca-67b5-4f73-9e3e-16bb989c7c5f/Kompromiteerimiseindikaatoritekasutaminekberin.pdf	“MITRE” and “Triage”	14.02.2023	Google Scholar
Using SLA Strategy to Design an SOC Platform in Data Center on the Cloud Computing	https://www.airitilibrary.com/Publication/alDetailedMesh?docid=16079264-201309-201310030016-201310030016-751-758	“SOC” and “Information Security”	13.02.2023	Google Scholar
VISNU: A Novel Visualization Methodology of Security Events Optimized for a Centralized SOC	https://ieeexplore.ieee.org/abstract/document/8453754	“SOC” and “Information Security”	13.02.2023	Google Scholar
Why Audit?	https://link.springer.com/chapter/10.1007/978-1-4842-2140-2_1	“SOC” and “IT Security”	16.02.2023	Google Scholar
Why IT Security Matters	https://link.springer.com/chapter/10.1007/978-1-4842-8628-9_1	“SOC” and “IT Security”	16.02.2023	Google Scholar
Why SIEM is Irreplaceable in a Secure IT Environment?	https://ieeexplore.ieee.org/document/8732173	“SOC” and “cybersecurity”	18.02.2023	IEE Xplore
Wide-area Cyber-security Analytics Solution for Critical Infrastructures	https://ieeexplore.ieee.org/document/9236483	“security operations center”	14.02.2023	IEEE Xplore

Appendix B

SLR Second Search Process Included and Excluded Sources

Title	URL	Search Phrase	Date	Database
A comparative study on cyber threat intelligence: the security incident response perspective	https://ieeexplore.ieee.org/abstract/document/9557787	cyber threat intelligence	04.02.2023	Google Scholar
A Cyber Security Data Triage Operation Retrieval System	https://www.researchgate.net/profile/Chen-Zhong-4/publication/323528335_A_cyber_security_data_triage_operation_retrieval_system/links/5f8f0a76458515b7cf90c7f5/A-cyber-security-data-triage-operation-retrieval-system.pdf	cyber incident triage systems	04.02.2023	Google Scholar
A Formal Approach to Analyzing Cyber-Forensics Evidence	https://www.webofscience.com/wos/woscc/full-record/WOS:000460205700014	cyber forensics	04.02.2023	Web of Science
A Framework for Cyber Threat Intelligence Extraction from Raw Log Data	https://ieeexplore.ieee.org/abstract/document/9006328	cyber threat intelligence	04.02.2023	Google Scholar
A Two-Step Approach to Optimal Selection of Alerts for Investigation in a CSOC	https://www.scopus.com/record/display.uri?eid=2-s2.0-85058624326&origin=resultslist&sort=r-f&src=s&st1=cyber+incident+triage+process&nlo=&nlr=&nls=&sid=5bcf8dddcd1d691f8721d3405de366f6&sot=b&sdt=b&sl=53&s=ALL%28cyber+incident+triage+process%29+AND+PUBYEAR+%3e+2012&relpos=107&citeCnt=2&searchTerm=	cyber incident triage process	04.02.2023	Scopus
Actionable Cyber Threat Intelligence for Automated Incident Response	https://www.scopus.com/record/display.uri?eid=2-s2.0-85147855038&origin=resultslist&sort=r-f&src=s&st1=cyber+threat+intelligence&nlo=&nlr=&nls=&sid=ae7aea43231119e1804980cdb049b621&sot=b&sdt=b&sl=49&s=ALL%28cyber+threat+intelligence%29+AND+PUBYEAR+%3e+2012&relpos=113&citeCnt=0&searchTerm=	cyber threat intelligence	04.02.2023	Scopus
Active Learning for Alert Triage	https://www.webofscience.com/wos/woscc/full-record/WOS:000353638700006	cyber incident triage process	04.02.2023	Web of Science
Addressing the increasing volume and variety of digital evidence using an ontology	https://www.scopus.com/record/display.uri?eid=2-s2.0-84920285002&origin=resultslist&sort=r-f&src=s&st1=cyber+incident+triage+process&nlo=&nlr=&nls=&sid=5bcf8dddcd1d691f8721d3405de366f6&sot=b&sdt=b&sl=53&s=ALL%28cyber+incident+triage+process%29+AND+PUBYEAR+%3e+2012&relpos=66&citeCnt=11&searchTerm=	cyber incident triage process	04.02.2023	Scopus
An Analysis of Digital Forensics in Cyber Security	https://link.springer.com/chapter/10.1007/978-981-13-1580-0_67	cyber forensics	04.02.2023	Google Scholar
Automated Cyber Threat Intelligence Reports Classification for Early Warning of Cyber Attacks in Next Generation SOC	https://www.webofscience.com/wos/woscc/full-record/WOS:000668350000009	cyber threat intelligence	04.02.2023	Web of Science
Automatic Narrative Summarization for Visualizing Cyber Security Logs and Incident Reports	https://www.scopus.com/record/display.uri?eid=2-s2.0-85123854170&origin=resultslist&sort=r-f&src=s&st1=cyber+AND+incident+AND+triage+AND+systems&nlo=&nlr=&nls=&sid=67bdb8812f9dd87cba64ec9d1452d518&sot=b&sdt=b&sl=65&s=ALL%28cyber+AND+incident+AND+triage+AND+systems%29+AND+PUBYEAR+%3e+2012&relpos=60&citeCnt=3&searchTerm=	cyber incident triage systems	04.02.2023	Scopus
Cyber Forensics and Comparative Analysis of Digital Forensic Investigation Frameworks	https://ieeexplore.ieee.org/document/9036214	cyber forensics	04.02.2023	IEEE Xplore
Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy	https://ieeexplore.ieee.org/abstract/document/7166125	cyber incident triage process	04.02.2023	Google Scholar
Cyber Security Threat Intelligence Monitoring and Classification	https://ieeexplore.ieee.org/document/9624746	cyber threat intelligence	04.02.2023	IEEE Xplore
CYBER THREAT MODELING FRAMEWORK	https://www.webofscience.com/wos/woscc/full-record/WOS:000848616300366	cyber threat intelligence	04.02.2023	Web of Science
Cyber Treat Intelligence Modeling	https://www.webofscience.com/wos/woscc/full-record/WOS:000490868600028	cyber threat intelligence	04.02.2023	Web of Science

Title	URL	Search Phrase	Date	Database
Detecting the software usage on a compromised system: A triage solution for digital forensics	https://www.scopus.com/record/display.uri?eid=2-s2.0-85143536196&origin=resultslist&sort=r-f&src=s&st1=cyber+incident+triage+process&nlo=&nlr=&nls=&sid=5bcf8dddcd1d691f8721d3405de366f6&sot=b&sdt=b&sl=53&s=ALL%28cyber+incident+triage+process%29+AND+PUBYEAR+%3e+2012&relpos=110&citeCnt=0&searchTerm=	cyber incident triage process	04.02.2023	Scopus
Discerning cyber threatening incidents from ordinary events using sentiment analysis and logistic regression	https://onlinelibrary.wiley.com/doi/full/10.1002/spy2.69	automation of cyber incident triage	04.02.2023	Google Scholar
Early Detection of Cybersecurity Threats Using Collaborative Cognition	https://ieeexplore.ieee.org/abstract/document/8537852	cyber threat intelligence	04.02.2023	Google Scholar
Finding cyber threats with ATT&CK-based analytics	https://apps.dtic.mil/sti/citations/trecms/AD1107945	cyber threat intelligence	04.02.2023	Google Scholar
Forensics for multi-stage cyber incidents: Survey and future directions	https://www.scopus.com/record/display.uri?eid=2-s2.0-85145262956&origin=resultslist&sort=r-f&src=s&st1=automation+of+cyber+incident+triage&sid=bae716038b5ef957953029e38a5a5315&sot=b&sdt=b&sl=78&s=ALL%28automation+of+cyber+incident+triage%29+AND+PUBYEAR+%3e+2012+AND+PUBYEAR+%3e+2012&relpos=31&citeCnt=0&searchTerm=	automation of cyber incident triage	04.02.2023	Scopus
Framework of Cyber Attack Attribution Based on Threat Intelligence	https://www.webofscience.com/wos/woscc/full-record/WOS:000406999100011	cyber threat intelligence	04.02.2023	Web of Science
Fronesis: Digital Forensics-Based Early Detection of Ongoing Cyber-Attacks	https://ieeexplore.ieee.org/document/10004506	cyber forensics	04.02.2023	IEEE Xplore
How Ready is Your Ready? Assessing the Usability of Incident Response Playbook Frameworks	https://dl.acm.org/doi/abs/10.1145/3491102.3517559	cyber incident triage process	04.02.2023	Google Scholar
Human factors in automating cyber operations	https://www.scopus.com/record/display.uri?eid=2-s2.0-85108336556&origin=resultslist&sort=r-f&src=s&st1=automation+of+cyber+incident+triage&sid=bae716038b5ef957953029e38a5a5315&sot=b&sdt=b&sl=78&s=ALL%28automation+of+cyber+incident+triage%29+AND+PUBYEAR+%3e+2012+AND+PUBYEAR+%3e+2012&relpos=19&citeCnt=1&searchTerm=	automation of cyber incident triage	04.02.2023	Scopus
Implications of Theoretic Derivations on Empirical Passive Measurements for Effective Cyber Threat Intelligence Generation	https://ieeexplore.ieee.org/document/8422720	cyber forensics	04.02.2023	IEEE Xplore
Improving Forensic Triage Efficiency through Cyber Threat Intelligence	https://www.mdpi.com/1999-5903/11/7/162	cyber incident triage process	04.02.2023	Google Scholar
Integrated Network and Security Operation Center: A Systematic Analysis	https://ieeexplore.ieee.org/abstract/document/9729852	cyber incident triage systems	04.02.2023	Google Scholar
Learning From Experts' Experience: Toward Automated Cyber Security Data Triage	https://ieeexplore.ieee.org/document/8360965	cyber incident triage process	04.02.2023	IEEE Xplore
Leveraging decision making in cyber security analysis through data cleaning	https://digitalscholarship.tsu.edu/sbaj/vol16/iss1/1/	cyber incident triage process	04.02.2023	Google Scholar
Security operations centre: Situation awareness, threat intelligence and cybercrime	https://ieeexplore.ieee.org/document/8073384	cyber threat intelligence	04.02.2023	IEEE Xplore
SOTER: A Playbook for Cybersecurity Incident Management	https://ieeexplore.ieee.org/abstract/document/9086465	cyber incident triage systems	04.02.2023	Google Scholar
Studying Analysts' Data Triage Operations in Cyber Defense Situational Analysis	https://link.springer.com/chapter/10.1007/978-3-319-61152-5_6	cyber incident triage process	04.02.2023	Google Scholar
Taxonomy of Cyber Threat Intelligence Framework	https://ieeexplore.ieee.org/document/9952616	cyber threat intelligence	04.02.2023	IEEE Xplore
The Adoption of Automation in Cyber Forensics	https://www.webofscience.com/wos/woscc/full-record/WOS:000848118200009	cyber forensics	04.02.2023	Web of Science
The Value of Metadata in Digital Forensics	https://ieeexplore.ieee.org/document/7379751	cyber forensics	04.02.2023	IEEE Xplore

Title	URL	Search Phrase	Date	Database
Threat intelligence: What it is, and how to use it effectively	https://nsfocusglobal.com/wp-content/uploads/2017/01/SANS_Whitepaper_Threat_Intelligence__What_It_Is__and_How_to_Use_It_Effectively.pdf	cyber threat intelligence	04.02.2023	Google Scholar
Towards an Automated Dissemination Process of Cyber Threat Intelligence Data using STIX	https://ieeexplore.ieee.org/document/9631850	cyber threat intelligence	04.02.2023	IEEE Xplore

Appendix C

Tested Tecnhiques

C.1 T1078 - Valid Accounts

Extracted Data		Alarm 1	Alarm 2	Alarm 3	Alarm 4	Alarm 5	Alarm 6	Alarm 7	Alarm 8	Alarm 9	Alarm 10
Sign in time	ok	Normal Time	Normal Time	Normal for US	Normal	Normal	Normal	Normal	Suspicious	Normal for US	Suspicious
Sign in location	ok	Suspicious	Normal	Normal for US	Normal	Normal	Suspicious	Normal	Normal	Normal	Normal
Sign in user agent	ok	Suspicious	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal
Success?	ok	no	no	Both	no	Yes	yes	Both	Yes	Both	Both
Sign in device	ok	N/A	Normal	ok	Normal	Normal	Normal	Normal	Normal	Normal	Normal
Failed authentication	ok	N/A	yes	no	N/A	no	no	yes	No	No	No
MFA	ok	no	Not Relevant	Not Relevant	Failed	Not Relevant	Success	Failed	Sucess	Failed	Sucess
VPN?		Not Relevant	Not Relevant	yes	Not Relevant	Not Relevant	Uncertain	Not Relevant	Uncertain	No	No
Abuseip?	ok	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Calendar?	ok	Not Relevant	Not Relevant	Not Relevant	Not Relevant	Not Relevant	Uncertain(!)	Not Relevant	Yes	Not Relevant	Not Relevant
Sign in anomalies				US login	MFA declined	Mobile network	Zurich		Login from Ireland	Login from US	Spike in failures(most likely expired token)

Upon the first iteration of valid accounts we reached a lower than expected success rate and had to perform a secondary discussion and testing phase. We discovered the issue to be the lack of result type in the log data, which subsequently resulted in us not being able to see if a login was performed or not. In order to know if valid accounts has happened, that is, an adversary has successfully logged in for example, it is necessary to see if a login failed or not. Information gathered in this step would also be beneficial to include when writing a report to the customer.

Result type	ok	Blocked	Expired refresh Token	KmsiInterupt, Strong authentication, expired token	MFA declined	Ok	Reauthentication	Wrong password/user name	Ok	KmsiInterupt	Expired refresh Token

C.2 T1204 - User Execution

Extracted Data	Alarm 1	Alarm 2	Alarm 3	Alarm 4	Alarm 5	Alarm 6	Alarm 7	Alarm 8	Alarm 9	Alarm 10
File Name	suspicious	normal	normal	suspicious	suspicious	suspicious	suspicious	normal	suspicious	suspicious
File Hash	malicious	ok	ok	malicious	no results	malicious	malicious	ok	malicious	malicious
Time	normal	normal	unnatural	unnatural	unnatural	unnatural	normal	normal	normal	normal
User	ok	ok	ok	ok	ok	unknown	unknown	ok	ok	ok
Download success	blocked	not relevant	not relevant	not relevant	yes	not relevant	not relevant	not relevant	yes	yes
Command run	not relevant	normal	not relevant	not relevant	none	unknown(?)	unknown(?)	normal	none	run installer
Network Traffic	not relevant	not relevant	not relevant	not relevant	unknown	not relevant	not relevant	not relevant	none	unknown
Device FIle Events	attempted dl	none	found by scan?	attempted install	found by scan?	found by scan?	found by scan?	seems update?	none	run installer
Device ID	ok	ok	ok	ok	unknown	ok	ok	ok	ok	ok
Hostname	ok	ok	ok	ok	unknown	ok	ok	ok	ok	ok

C.3 T1566 - Phishing

Extracted Data	Alarm 1	Alarm 2	Alarm 3	Alarm 4	Alarm 5	Alarm 6	Alarm 7	Alarm 8	Alarm 9	Alarm 10
Check URLs/fileHash	malicious	malicious	malicious	malicious	malicious	malicious	malicious	malicious	malicious	malicious
Check Network Traffic	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Check header info	Not Relevant	N/A	N/A	Not Relevant	N/A	N/A	N/A	N/A	N/A	N/A
Check other recipients	N/A	yes(but blocked	no	N/A	yes	no	N/A	N/A	N/A	N/A
Check if ran file/attatchment	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Phishing turned out to be quite a difficult technique to test, given the issue that there were several pieces of relevant information not present in the already extracted data presented in the SOAR overview. Building on this, there were information we either did not have the access to extract or the knowledge of how, and learning would leave the risk of taking too much time, which we did not have. As a solution we elected to ask the supervisor, if given the log data as presented in the table, would we be able to reach a conclusion on the alarms?

The drawn conclusion on phishing was that we had a pretty accurate solution in the beginning, much thanks to our supervisor participating in the discussion phase and providing input. We did however want to include a check on the senders reputation as a metric to judge the legitimacy of a received e-mail. The problem we discovered that posed, was that it is very difficult to check if an address is legitimate, seeing as addresses used in phishing can often be automatically generated or new, and as such have no reputation.

When discussing with our supervisor he also pointed out some other log sources such as commonsecurity logs(firewall logs) and advanced hunting, which would give us the ability to detect if the user have had any communication with the

sender(clicked the link or attachment) or any files have been downloaded and/or executed, created, modified etc. Another useful log source he mentioned was the Microsoft purview which could be used to look up if there were any other recipient of the same potential phishing e-mail.

Bibliography

[1] "Alexander Maedche""Jan vom Brocke" "Alan Hevner". *Introduction to Design Science Research*. 2020. URL: https://www.researchgate.net/publication/345430098_Introduction_to_Design_Science_Research. (accessed: 10.04.2023).

[2] "Salvatore T. March" "Jinsoo Park" "Sudha Ram" "Alan R. Hevner". *THE DESIGN SCIENCE RESEARCH PROCESS: A MODEL FOR PRODUCING AND PRESENTING INFORMATION SYSTEMS RESEARCH*. 2004. URL: https://www.jstor.org/stable/25148625. (accessed: 10.04.2023).

[3] "David Jaeger""Feng Cheng""Christoph Meinel" "Amir Azodi". *Pushing the Limits in Event Normalisation to Improve Attack Detection in IDS/SIEM Systems*. 2013. URL: https://ieeexplore.ieee.org/document/6824575. (accessed: 22.03.2023).

[4] "Dimitris Askounis" "Anna Georgiadou" "Spiros Mouzakitis". *Assessing MITRE ATT&CK Risk Using a Cyber-Security Culture Framework*. 2021. URL: https://www.mdpi.com/1424-8220/21/9/3267. (accessed: 14.02.2023).

[5] "Stuart M. Charters" "Barbara Kitchenham". *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007. URL: https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering. (accessed: 02.02.2023).

[6] "Hossein Saiedian" "Blake D. Bryant". *Improving SIEM alert metadata aggregation with a novel kill-chain based classification model*. 2020. URL: https://www.sciencedirect.com/science/article/pii/S016740482030095X. (accessed: 14.02.2023).

[7] "IBM". *What is SIEM? Security information and event management (SIEM) explained*. URL: https://www.ibm.com/topics/siem. (accessed: 22.03.2023).

[8] "Tuure Tuunanen" "Charles E. Gengler" "Matti Rossi" "Wendy Hui" "Ville Virtanen" "Johanna Bragge" "Ken Peffers". *THE DESIGN SCIENCE RESEARCH PROCESS: A MODEL FOR PRODUCING AND PRESENTING INFORMATION SYSTEMS RESEARCH*. 2006. URL: https://arxiv.org/ftp/arxiv/papers/2006/2006.02763.pdf. (accessed: 10.04.2023).

[9] "Ronald Paans" "Stef Schinagl" "Keith Schoon". *A Framework for Designing a Security Operations Centre (SOC)*. 2015. URL: https://ieeexplore.ieee.org/document/7070084. (accessed: 14.02.2023).

[10] "Maria Watson" "Yu Xiao". *Guidance on Conducting a Systematic Literature Review*. 2019. URL: https://journals.sagepub.com/doi/full/10.1177/0739456X17723971. (accessed: 02.02.2023).

[11] MITRE Corporation. *ATT&CK Matrix for Enterprise*. 2013. URL: https://attack.mitre.org/. (accessed: 14.02.2023).

[12] "MITRE CORP ANNAPOLIS JUNCTION MD". *Finding Cyber Threats with ATT&CK™-Based Analytics*. 2017. URL: https://apps.dtic.mil/sti/citations/trecms/AD1107945. (accessed: 04.02.2023).

[13] "Palo Alto Networks". *What is SOAR?* URL: https://www.paloaltonetworks.com/cyberpedia/what-is-soar. (accessed: 22.03.2023).

[14] "Cyril Onwubiko". *Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy*. 2015. URL: https://ieeexplore.ieee.org/abstract/document/7166125. (accessed: 14.02.2023).

[15] "MITRE CORP McLean VA". *MITRE ATT&CK®: Design and Philosophy*. 2018. URL: https://apps.dtic.mil/sti/citations/trecms/AD1107945. (accessed: 29.03.2023).