UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering, Mathematics & Computer Science

A Study on Blue Team's OPSEC Failures

Matthias Caretta Crichlow
October 2020

NORTHWAVE
Intelligent Security Operations

Supervisors:
dr. Jeroen van der Ham
Bart Roos

Acknowledgements

I would like to thank all the Team of Northwave for the opportunity to work in such an exciting environment. Special thanks to Marvin and all the analysts in Northwave's SOC for participating with enthusiasm in the research. To all the folks of the Red Team for being part of this journey. Bart Roos for his valuable advice and for helping me find the right direction for my research. dr. Jeroen van der Ham for his continued guidance and feedback throughout the research.

I dedicate this work to all the members of my family for believing in me and always supporting my studies. And to my Mom and Dad who taught me to work hard, always think outside the box and motivated me to give 1000% in every situation.

Abstract

Organizations are every day expanding their networks, increasing the number of servers and workstations in it. Such a growth expands the surface that can be targeted by malicious actors to cause harm. Therefore it is becoming more and more common for the organizations to create specialized teams of defenders (i.e. the Blue Team) who can monitor and protect their system. However, the fact that someone is actively hunting for malicious actors changed the balance in cybersecurity. Interacting with the attackers causes change in their strategies. We focused our efforts in studying the interplay between attackers and defenders, aiming at creating further studies in this new field. As the first step we tried to understand what part of the Blue Team investigations can be detected by an intruder, and we highlighted the fact that indicators of Blue Team's OPSEC failures are the way attackers can likely achieve these results. We focused our study on the first line of defence within the Blue Team, the SOC (Security Operation Center). Using CTA (Cognitive Task Analysis) techniques we identified common OPSEC failures among SOC analysts. Subsequently, in order to evaluate the impact that such actions have on the strategies of attackers we organized a wargame in collaboration with Northwave's Red Team demonstrating that being aware of the Blue Team's presence determined the adoption of more cautious behaviour in the attacker. In order to achieve our goal we developed a new CTA technique that can be used to further study Blue Team's cognitive processes. Additionally, we addressed a major problem within the cybersecurity research community by developing a reusable virtual environment with built-in monitoring capabilities that can be used to create experiments that can be easily verified by other researchers.

Contents

Chapter 1

Introduction

This research aims to improve the general understanding of the interplay between the offensive and defensive actors in the cybersecurity realm. More specifically, the goal is to study the dynamics between the Red Team and Blue Team during a Red Team assessment.

1.1 Problem Statement

A Red Team exercise is a tool used to test an organization IT infrastructure in a realistic cyber-attack scenario. The Red Team (RT) emulates the actions of an adversary and tries to breach into the organization's network. If the organization is mature enough to have an active Blue Team (BT), they will try to respond to the attack in real-time and to take appropriate countermeasures such as shutting down or isolating infected endpoints, deleting malicious files or stopping harmful processes.

The issue that any Red Teamer (but also an attacker in general) faces is that of unbound uncertainty. When performing an attack, there are two possible outcomes: success or failure. However, troubleshooting the reasons for failure is often just guesswork. Questions like: Was my malware delivered? Why was it not delivered? Was it executed? Why was it not executed? Is it being run in a sandbox? And several other questions like these currently remain unanswered. There are too many variables out of Red Team's control, and in many situations, the RT ends up operating blindly and making uninformed decisions. This, in turn, leads to a waste of time, resources and opportunities. Understanding the reasons for failures is crucial

to make informed decisions on the next steps and maximize the probability of successfully breaking into a system. The problem is that a feedback loop for failures in offensive cyber operations does not exist. This issue becomes even more relevant when a Blue Team is actively investigating the cyber attack. In order to improve the effectiveness of the Red Team it is necessary to better understand the relationship between the two teams.

When the Blue Team is actively investigating traces of cyber attacks, the Red Team is faced with an adversary that is not only capable of responding to single attacks but with an adversary that is capable of put together the pieces of the puzzle and stop a whole offensive campaign. The presence of a human actor in the game contributes even more to the uncertainty of the Red Team. A Blue Team investigation can indeed be seen as an inconvenience, but also as an opportunity. It would be difficult to extract information from a system that is isolated from the outside world. According to Locard's exchange principle[1]:" Every contact leaves a trace", which means that while the Blue Team investigates traces of cyber attack, they will inexorably give away something about their operations. This is often referred to as an *OPSEC failure*. OPSEC failures of the Blue Team create windows of opportunity that an attacker can use to infer the reasons why their attack was blocked and even more by giving the attacker insights into Blue Team operations.

Understand which actions of the Blue Team are detectable by an attacker is very important, not only to help the Blue Team to perform more secure investigations but also for the Red Team to increase the success rate of their assessments. However, there is a lack of scientific publications which study the point of view of the attacker. More specifically, the influence that a Blue Team investigation has on the behaviour and strategies of the attacker has not been addressed before. Fill such a research gap is the primary driver of this research.

SOC analysts For this research an investigation will be defined as all the actions done by the Blue Team from the moment a suspicious event is detected to the moment they decide to take appropriate action to respond to the threat. The Blue Team is composed of many different Teams, each one contributing to the security posture of a company differently. However, among all the various teams the SOC (Security Operations Center) has been chosen as the main subject of this research, the rea-

[1] Dr. Edmond Locard was a pioneer in forensic science. In forensic science, Locard's principle holds that the perpetrator of a crime will bring something into the crime scene and leave with something from it, and that both can be used as forensic evidence. [wikipedia]

son of this decision will be explained in more details in section 2.2. The complete spectrum of the interactions between attack and defender is very wide and complex; for this reason, this research can be considered a starting point in the study of Red and Blue Team interaction and hence address only the initial part of this interaction. Among the various teams the SOC team (Security Operations Center) is the first line of defence in an organisation, they collect and analyse security events and advice on which action should be taken next. This is the main reason why the SOC has been selected amongst the other components of the Blue Team as the subject of this research.

There are many more factors that influence the interaction between attacker and defenders. For example, some of those factors are: the Blue Time time to response; the actions taken by the Blue Team to stop an attacker from progressing further in the network; the motives of the attacker; the specific network topology; the technologies used to detect and respond to threat and many more. However, only the OPSEC failures of the Blue Team will be considered in the scope of this research. The reason for this decision is that OPSEC failures are identified as the most effective factor an attacker can exploit to gain an advantage on the Blue Team. The main research question is therefore defined as: *RQ1* - **How the Red Team can detect SOC analysts OpSec failures?**.

1.2 Research Questions

Given the nature of the problem under analysis, it is not possible to fully answer the research question by studying only one of the two main actors (Red and Blue Team), as a matter of fact, actions of the one influence the action of the other and vice-versa. For this reason, the structure of this research will reflect the duality of the problem. The main research questions will be broken down in two macro research questions; each one focused on a different actor. The result will then be combined to answer the main research question.

Blue Team research question The first research question aims at understanding which are the traces that the Blue Team might leave behind during their investigations. As it will be discussed in chapter 2, among the existing literature analysed none has been found which describes the footprints of Blue Team on a system. Before been able to answer the question ”how to detect something” it is necessary to

be able to answer to ”what can be detected”. For this reason, the first sub research question tries to discover what are the elements of a Blue Team investigation that an attacker can see. More specifically, which OPSEC failures a SOC analyst might do when investigating a security event. The first sub research question is, therefore:*RQ2* - **Which are the most common OPSEC failures among SOC analysts?**

Red Team research question The second part of this research aims to understand what is the actual impact of the Blue Team actions on the activities and strategies of the Red Team. There is an obvious difference between attacking a system which is protected and monitored by a defence team and a system which is not. However, it is not clear if this difference has any weight on the decisions of the attacker or not. The second sub research question is, therefore:*RQ3* - **To what extent a SOC analyst investigation influences the actions of the Red Team?**.

Ground for future research Being this research one of the first efforts in the direction of a better understanding on the Red and Blue Team interplay an additional goal is to lay a solid ground for future researchers who might want to further research this topic or to verify the result of this research. For this reason, besides answering the research questions, we aim at developing an easily reproducible testing environment to study the interaction between the two teams.

Chapter 2

Literature Review

2.1 Red Team

Cyber Red Teaming is a young concept and therefore at the time of writing, there are very few publications which provide a complete description of what Red Teaming is. This section has two goals, the first one is to introduce the reader to what is the Red Team and how it operates, and the second one is to provide a complete overview on the discipline of Red Teaming for reference of future researchers.

2.1.1 Adversarial Thinking

Red Team is a wide discipline that applies to many fields: intelligence, business, national security and cyber security. The very core idea is to look at a problem form an adversary or competitor perspective and provide the decision makers with the necessary information to take a weighted decision [1]. In a broad sense RT is used to reduce the impact of cognitive biases such as group think or confirmation bias. These biases arise when people are faced with too much information and use cognitive shortcuts to reach to conclusion fast [2]. RT uses a class of techniques called alternative analysis to challenge conventional thinking and force the organization to explore unconventional paths. This idea of RT is typically used to support decision making process in military or business field.

Group think To understand the Red Team is important to first understand the reasoning biases that the adversarial thinking technique tries to solve. Irvin Janis [3]

defines Groupthink as the tendency of group members to value the group higher than anything, to the point that they strive for reaching a painless unanimity on the issue the group has to confront [4]. Consequence of groupthink is that there is a lack of creativity in the proposed solution, and this often leads to suboptimal decision due to the lack of opposition. Groupthink brings some benefits, such as faster convergence to a decision and less conflicts, but also inhibits the ability of the group to see the bigger picture, group members do not raise objections or ask critical questions that would otherwise be overseen.

Confirmation bias Another common cognitive bias is the confirmation bias. It is the tendency of people to see evidence consistent with their pre-existing beliefs, in such a way they cannot see their own mistakes and consistently overlook some evidence or overestimate others. People that are victims of confirmation bias are focusing on one possibility and ignoring alternatives, this may also lead to overconfidence [5].

2.1.2 Red Teaming vs Pentesting

In the cyber domain, Red Teaming is still a young discipline and therefore is not yet well defined and is often confused with two other practices used to improve cybersecurity, i.e. penetration testing and vulnerability assessment.

Red Team (RT) and Pentesting (PT) are two ways to improve cyber defences. Both use similar tools to perform cyber attacks, but they differ in terms of goals and results. Red Teaming is focused on the "depth" of the assessment, while the pentest is aimed at covering the largest number of attack vectors – covering the "width" [6]. The following section will highlight differences and similarities between the two.

Penetration Testing is a type of security assessment conducted on information systems to identify vulnerabilities that could be exploited by cyber attackers [7]. During a Penetration test the assessors (often referred to as ethical hackers) use techniques and tools to duplicate the steps of cyber attackers when they try to breach into a system. The ethical hackers mimic the attacker only on a technical level. The test can be conducted on hardware, software or firmware components trying to find working exploits to bypass the defence mechanisms protecting such components. However, not every component in the system is the target of the pentester, and what can be tested is specifically defined by the scope of the assessment before

commencing it [8], [9]. The scope can be as specific as testing a certain web application, or as wide as testing the whole organization according to the needs of the customer. While the scope defines *what* can be tested the Rules of Engagement (RoE) *how* the testing can be conducted. RoE, for instance, may include the time of the day to test (to avoid business hours), and how sensitive data should be handled. The RoE can also include the locations the pentester may need to travel to in order to perform the test.

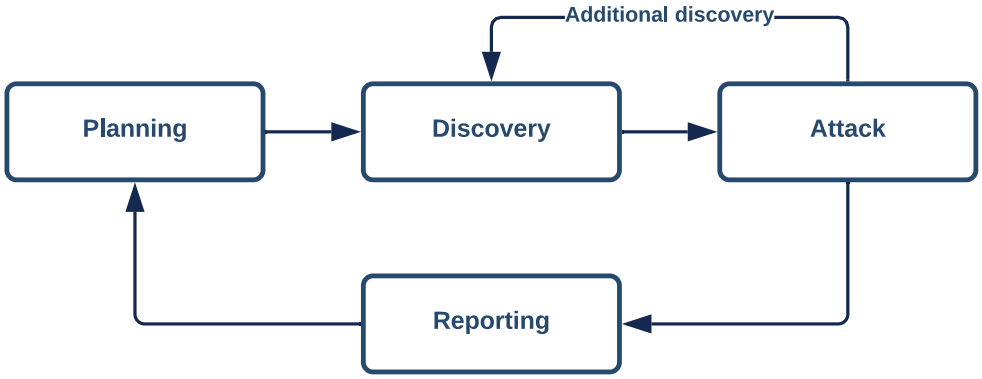


Figure 2.1: NIST 4-phases pentest

The NIST describes PT as a four-stage process [9] (See fig. 2.1). The first phase, planning, involves the steps described above about scoping and RoE. The discovery phase is divided into two steps: the first covers information gathering and scanning of the system. In the second step, the results are compared against vulnerability databases and combined with the tester knowledge about vulnerabilities. The next step is the actual attack. In the attack phase, the pentester attempts to exploit the discovered vulnerabilities. If the attempts are successful, the tester can try to escalate privileges. In this way, gaining more knowledge about the system and perform the discovery phase again with the newly acquired clearance level to find and exploit even more vulnerabilities. The conclusive phase is reporting. In this phase the pentester develops a report containing the identified vulnerabilities and suggestions to mitigate them.

Finally, it is worth mentioning that a pentest is just a resemblance of a real attack because the test is conducted within a set of constraints such as time, resources and the skills of the pentester. The outcome is, therefore, more valuable the more capable and knowledgeable the pentester is. Another factor that distances the PT from a real attack is the amount of information given to the pentester. Based on that the test can be divided into three types: black box, grey box and white box [10]. A black box is the test type in which no information at all is disclosed to the pentester.

In a grey box kind of test, some info is given to pentester such as network topology or the credentials of some low privileged users. A white box is the test type where the pentester has full knowledge about the target system/systems (See fig. 2.2).

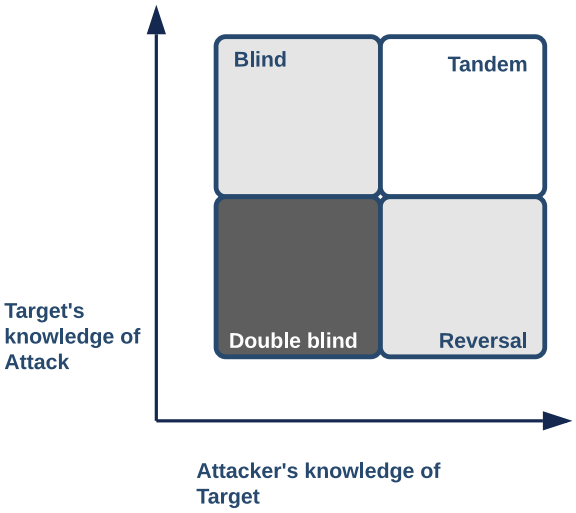


Figure 2.2: Testing types

Red Teaming is a broad discipline that applies to several domains such as business, military and cyber to support decision-making. The correct way to address RT in the cyber realm should be Cyber Red Teaming (CRT). However, from now on the terms Cyber Red Teaming and Red Team will be used interchangeably; that is because the focus of this work is limited to study the Red Team solely in the cyber domain. It is a common opinion amongst cybersecurity expert that there is a lack of clarity on the definition of Cyber Red Teaming , and so it is often confused with the terms 'penetration testing' and 'vulnerability assessment' [11], [12].

H. Dalziel [12] gives a simple yet clear explanation of the difference between Red Teaming and Penetesting: Cyber Red Teaming is goal-based, whereas PT and vulnerability assessment are target-based. What this means is that PT has a target which for instance can be: a web application, a server or a group of employees to do social engineering on. Then they may focus on that target and try to find and exploit as many vulnerabilities as possible. Vice versa RT sets an high level goal at the beginning of the assessment, which can be, for instance, to compromise customer data , find ways to get into the internal network, and compromise a certain business critical process. Once the goal is set, each action the Red Teamer takes should aim at taking him one step closer to achieve it, just in the same way a real attacker would do it [12]. A real attacker would not limit himself to attack just a specific system or to

use a specific set of technologies, but would instead use his creativity and combine different Techniques Tactics and Procedures(TTP) to achieve his goal. RT adopts a holistic approach to cybersecurity. It incorporates different organization's elements in the assessment, such as network systems and software, as well as business processes. An example of a business process incorporated in a RT assessment is exploiting the organization's hiring procedure to get physical access to facilities and establish an initial foothold.

A Red Team exercise consists of simulating adversarial attempts to compromise organizational Critical Functions(CF) [1] and the information system supporting such functions. Just like real attacks the simulated attacks can target the *technology* (e.g., interactions with hardware, software, or firmware components) as well as the *people* (e.g., interactions via email, telephone, shoulder surfing, or personal conversation), and *physical facilities* (e.g., locks, physical access to a network, dumpster diving, intrusion testing) [7]. The goal of RT assessment is to perform a controlled and realistic cyber-attack simulation against an organization to test its detection and response capabilities. However, make a Red Team exercise that resembles a real-life attach requires a thorough intelligence work that gathers knowledge about the adversary's techniques, mindsets and goals [11]. Emulating an attacker allows an organization to have at its disposal an actor that thinks "outside the box". Such an actor can spot vulnerabilities and weaknesses that who planned the defences might not have foreseen.

2.1.3 Adversary Modelling

The main requirement to perform a RT assessment is being able to anticipate and replicate adversarial behaviour [11]. Therefore it is important to have a deep understanding of the adversary, and to have a disposal of a set of frameworks that can be used to model a malicious actor. This section will first provide some background knowledge about the motives of cyber attackers. Then it will introduce the most prominent frameworks used to model attackers capabilities and modus operandi [2].

[1] Critical Functions are business functions or services that if compromised, would significantly impact business continuity [13]

[2] Modus operandi (often shortened to M.O.) is someone's habits of working, particularly in the context of business or criminal investigations [14]

2.1.4 Understanding the attackers

Adversary motivation In their yearly report on threat actors the RAND corporation[3] discussed the motivations of the various types of malicious groups [15]. The author argues that cyber threat actors can be grouped based on their goals, motivations and capabilities. Based on those factors, four categories are suggested: Cyber terrorists, hacktivists, state-sponsored actors and cybercriminals. Robinson et al. [16] argued that there are three more categories that should be included in this list: script-kiddies, cyber researchers and internal actors. In the context of a Red Team assessment, the last three types are not of particular interest. Script kiddies do not have enough skills to be a severe threat to an organization, and cyber researchers are not motivated by malicious intentions.

Types of cyber criminals *Cyberterrorism* is the act of conducting terroristic attacks through cyberspace, intending to cause severe harm or death. There are currently no real-world examples of cyberterrorism [4]. However, we can expect to witness cyberterrorist attacks in the future, due to the increased integration between cyber and physical world. *Hacktivists* are instead motivated by an ideology or by a cause (political, social or economic). Unlike cyberterrorism, the aim is to expose information or disrupt a system, but not to cause any harm to people. *State-sponsored* actors motivation is to advance the interests of the nation-state that is funding them. They are also the most sophisticated in this list and able to perform long and complex attacks. Finally, *cybercriminals* are motivated by financial gain; they will try to acquire valuable information and then to sell them on the underground market.

2.1.5 Frameworks

Cyber Kill Chain

In 2011 Lockheed Martin developed a model called Cyber Kill Chain that expands the traditional military F2T2EA[5] chain model into one specific for cyber intrusions. The Cyber Kill Chain also known as Intrusion Kill Chain is defined as a series of six

[3]The RAND Corporation is a research organization that develops solutions to public policy challenges

[4]arguably STUXNET [17] can be considered an example of cyberterrorism

[5]U.S. Department of Defence describes it as an integrated end-to-end process divided into six steps: Find, Fix, Track, Target, Engage, Assess.

steps. Each step comes after the previous one without exception, it is described as ”chain” because any defincency will interrupt the entire process [18]. The elements of the chain are: Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control (C2C), Actions on Objectives (See fig. 2.3).

- **Reconnaissance** consists in gathering information about the target, identifying and profiling it. This step can be further broken down into passive and active reconnaissance. *Passive reconnaissance* is carried out by collecting information without directly interacting with the target.*Active reconnaissance* requires a more deep profiling of the target by directly interacting with it, and this may raise alarms.
- **Weaponization**, At this stage the attacker uses the vulnerabilities and the knowledge about the target acquired in the previous phase to craft malware that can exploit them.
- **Delivery** This stage involves transmitting the weapon to the target of the attack. Accomplish this task often may require the attacker to be creative and use social engineering techniques, as well as delivering usb drives containing the weapon.
- **Exploitation** Once the weapon is delivered to the victim the malicious code is triggered. The triggering can happen through remote or local mechanism (e.g. actions of the victim)
- **Installation**, at this step the malware installs backdoors, downloading additional software in order to allow the attacker to maintain persistence inside the environment.
- **Command and Control (C2)** this phase starts once the attacker has a communication channel with the compromised target inside the network. This way the attacker can send remote instructions to the compromised machines.
- **Actions on Objectives** is the final stage. An intruder can from now on take actions to achieve their original objectives. The command the attacker will execute depends on his intentions; it is possible to exfiltrate data but also to use the compromised system as a hop point to hack additional systems in the network performing lateral movement.

Use of Cyber Kill Chain The Cyber Kill Chain (CKC) is a useful tool to support defences when used to analyze intrusions post-mortem. After an incident has occurred and has been detected an analyst can go backward through the steps that

Figure 2.3: Cyber Kill Chain

lead the attacker inside the network. In this way it is possible to reconstruct the events and have a better understanding of what went wrong. Moreover it is possible to strategically compare multiple intrusions over time, identify commonalities and correlate indicators allowing the analyst to link together activities from the same threat actors and discover bigger campaigns [18]. Discovering patterns and behaviours can help the analysts to understand an intruder's intents and objectives. Hence, planning focused security measures to better defend the targets of such campaigns.

Criticisms to CKC R. Stolte sustains that the classic kill chain model was designed to fight against external threats, but many people wrongly try to used the CKC to model other kinds of threats, such as insiders threats [19]. Insider threats have a different behaviour from outsiders. Many of today's threats did not exist when the CKC was first conceptualized. It is not a criticism of the CKC itself but of the faulty way the model is used to model certain actors. Another criticism to CKC is that it reinforces old-school, perimeter focused, malware-prevention thinking [20]. The author sustains that modern threats thrive between the phases command&control and action on objective. However, the CKC fails to capture their behaviour between these two phase.

Unified Cyber Kill Chain

To solve the limitation of the CKC and improved version has been developed. Paul Pols argued that the CKC is limited to modelling the initial compromise of the system. The Unified Cyber Kill Chain (UKC) is a model that covers the attack phases that occur behind the organization's perimeter. It improves the CKC because the UKC phases may be bypassed, occur more than once, or out of sequence [21]. The main difference is that stopping an attacker at any phase of the sequence is no more enough to disrupt the whole chain, as an attacker can easily dodge countermeasures and move to a different stage. The Unified Cyber Kill Chain stimulates the deployment of a layered defence strategies and defense in depth principles (See fig. 2.4).

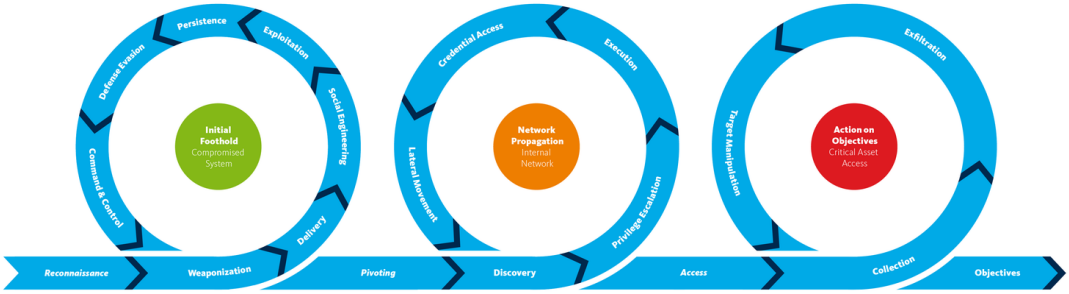


Figure 2.4: The Unified Kill Chain

ATT&CK Framework

MITRE [6] Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) is a knowledgebase for cyber adversary behavior. ATT&CK MITRE systematically analyzes and categorizes the TTPs adversary beahaviours and serves as both a model and a framework. The framework aims at improving the ability of detecting post-compromise adversary actions. It is meant to advance cyber threat intelligence (CTI) by establishing a generic vocabulary to describe post-compromise adversary behavior [22]. However, another version of the framework has been recently released called PRE-ATT&CK. The new ”flavor” of the framework covers the actions and the goals of an attacker before entering an organization’s network.

This framework is a collection of Tactics Techniques and Procedures observed in real Advanced Persistent Threats (APT). For this reason it can serve both offensive and defensive purposes. It can be used as an adversary emulation playbook that, for instance, a Red Team may it to developing realistic scenarios and to emulate adversary. But it can also be used as method for discovering defence gaps inside a network [23]. It is divided into tactics and techniques[7]. *Tactics* are high-level goals an attacker has during an operation, and they describe **why** an adversary perfrom a certain action. *Techniques* are the actions an adversary take to achieve the tactical objectives. Techniques describe **how** the attacker can act to accomplish his goals [23], [24].

The existing models such as Cyber Kill Chain, the Unified Cyber Kill Chain etc.

[6] MITRE is not an acronym, although many mistakenly believe it stands for Massachusetts Institute of Technology Research & Engineering

[7] Conversely to Indicator of Compromise (IoC) which look at the results of an attack, Tactics and Techniques are a way to look for on-going attacks

describe at high level the processes and adversary goals, but they are not adequate to describe what actions attackers make. On the other hand, low level sources of information such as malware databases and exploit databases contain information of specific instances of software and do not provide context around those information. The ATT&CK framework is mid-level software that allows to put low level concepts into context [24]

2.1.6 Red Teaming standards

Modelling the adversary is just one part of a more complicated process, which is the RT assessment. Being Red Team assessment a relatively new discipline it still lacks a unique definition of how it should be performed. There are various standards and guidelines across the globe that try to define what are the main elements of a Red Team assessment. The following section will compare the main standards in order to give a clear definition of the process followed during a Red Team assessment. Sub research question II will require an in-depth understanding of the process which guides the actions of the Red Team. This section will provide such knowledge in order to contextualise the activities of the Red Team.

Due to the fact that the actions of adversary groups have been particularly aggressive toward some specific industry sectors, penetration testing and red teaming assessment are often required by governments and certification authorities. A report from Boston Consulting Group [25] shows that the financial sector is more than 300 times more likely to be the target of cybercriminals; the reason for that is that the motivation of this kind of attackers is financial gain. At the same time, banks and financial institutions are more inclined to invest more in cybersecurity to protect their assets. Therefore, it is not a surprise that most of the guidelines come from financial institutions. CBEST, TIBER-NL, TIBER-EU, iCAST, AASE and FEER are all cyber-attack simulation frameworks developed by financial institutions to define how a Red Team exercise should be performed, and what are the prerequisite and desired outcome for the exercise.

CBEST The bank of England launched the CBEST framework in 2014. A couple of spinoff namely GBEST and TBEST where then proposed to address the needs of the government and telecommunication industry. This framework is used to test how much an organization is susceptible to cyber attacks. CBEST is a framework

that supports intelligence-led penetration testing operations to mimic the action of cyber attackers. The assessment process described by CBEST is composed of four phases: Initiation phase, Threat intelligence phase, Testing, Closure phase [26].

TIBER The Dutch National Bank (DNB) created Threat Intelligence-based Ethical Red Teaming (TIBER-NL) to increase the resiliency of Dutch financial institutions to cyber attacks. They describe the test as "the highest possible level of intelligence-based Red Teaming exercise using the same Tactics, Techniques and Procedures (TTPs) as real adversaries, against live critical production infrastructure, without the foreknowledge of the organisation's defending Blue Team (BT)". Actually a small group of people from the organization knows about the test, they are called the White team. The TIBER-NL framework was subsequently used by the European Central Bank (CBE) to create an European version of TIBER (TIBER-EU) . The EU aimed at creating a framework that could then be redefined and adopted by other jurisdictions as well. So far the framework has been adopted in Belgium, Denmark, Sweden, Germany and Ireland. There are no major differences in the aforementioned implementations of TIBER-EU, but each jurisdiction can adapt the framework in a manner that suits its specificities. The main advantage of adopting a common standard across countries in Europe is that it eases cross jurisdictional testing of organizations that are active in more than one country. TIBER divides the Red Teaming process in four phases: Generic Threat Intelligence phase, Preparation phase, Testing phase, Closure phase. Considering the criticality of the systems under test it is possible to cause damage to critical live production systems or even to lose or compromise sensitive data. Therefore TIBER advise to perform risk assessment on the risk posed by Red Team assessment itself, and requires the planning of escalation procedures in case of incident. [27].

iCAST Intelligence-led Cyber Attack Simulation Testing is a framework introduced by the Hong Kong Monetary Authority (HKMA). It augments the traditional Penetration testing introducing threat intelligence elements to create real life testing scenarios. The process defined in iCAST is divided in three main phases: Initiation, Intelligence gathering, Testing [28].

AASE Adversarial Attack Simulation Exercises is a framework developed by the Association of Banks in Singapore(ABS) to challenge the security defenses of an organization by targeting it with attacks based on real adversary techniques. The

stated goal is to provide the organization with insight on weaknesses that might not be found by standard security assessment methodologies such as vulnerability assessment and penetration testing. Similarly to the frameworks discussed above AASE is composed of four phases: Planning, Attack Preparation, Attack Execution, Closure [13].

FEER The Financial Entities Ethical Red Teaming framework was developed by the Saudi Arabian Monetary Authority as a guide to prepare and execute controlled attacks against live production environments. Unique feature of this framework is that it includes the use of a Green Team together with the canonical Red, Blue and White teams. The green Team represents the Financial supervisory authority whose role is to guide and support the white team during the exercise. The framework consists of four phases:Preparation phase, Scenario elaboration phase, Execution phase, Lesson learned phase [29].

Based on the presented frameworks, it is possible to see the common pattern described by the authors in the FSI's report [30] on the way the tests are performed. An initial phase to define the scope, the critical functions and the assets supporting these functions (see figure 2.5. An intelligence phase where the relevant threat actors are identified and modelled listing the TTPs they use. A scenario phase where the intelligence gathered during the previous phase is used to define attack scenarios that will lead the test phase. A testing phase during which the Red Teamers perform the actual test targeting the people, the processes and the system supporting the critical functions. And finally, a closure phase that involves the collaboration of the blue and Red Team to perform replay exercises, and the sharing of the lesson learned with other organizations of the sector.

2.1.7 Conclusion

In conclusion, the previous section described what is Red Teaming and showed that Red Teaming is more than just a technical tool. Indeed it can be applied to multiple different scenarios to brake out of cognitive bias loops. When applied to the technical field, it is better referred to as cyber Red Teaming. However, it is essential to remember that cyber Red Teaming is not just a technical operation (like pentesting); it targets an organization at 360°, including people and process as well. Moreover, it was presented how the Red Team emulate the way the attackers think and the frameworks which are used to model their actions. Finally, this section identified

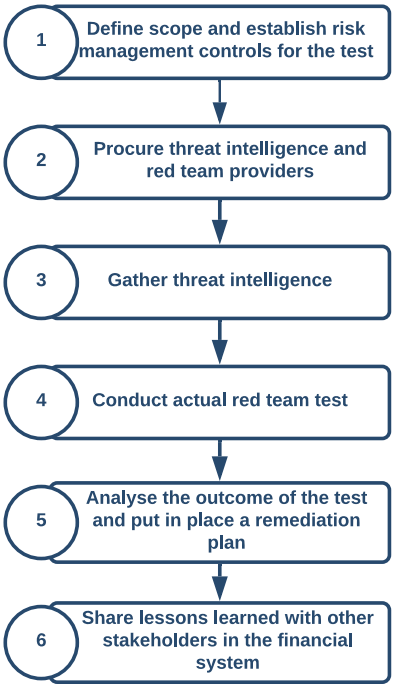


Figure 2.5: Highlevel Red Team Assessment Process described in [30]

what are the more common steps of a Red Team assessment by presenting various international standards for Red Teaming.

2.2 Securiy Operation Center

The following section will give an overview of the Blue Team operations, and examine what the state-of-the-art of the research on Blue team is. First, the relevant frameworks for Blue Team operations are analyzed. The frameworks define the background to understand the RT operation. Then the components of the Blue team are discussed in order to further narrow down the fundamental research question. Finally, it is important to understand how the investigative process works and how other researchers have approached similar problems. Therefore the relevant literature on the topic ”cyber attacker and defender interplay” will be presented.

2.2.1 Frameworks

There are a number of frameworks that supports cyber defence operations; some of them are a collection of elements; others describe processes. The following sec-

tion highlights the most relevant cybersecurity framework that directly supports Blue Team operations. It is important to mention that the following frameworks are the defensive-focused frameworks. However, the Blue Team often use also offensive-focused frameworks, such as the ones discussed in section 2.1.5.

The OODA loop is a framework developed by the U.S. Air Force to support fast decision-making process. Nowadays it is applied to many different fields (business, law enforcement, military and cybersecurity) and there exist many variants of it. It is a four-step cyclic process composed of: observer, orient, decide, act. The first stage, *observe*, aims at gathering information about the environment and the adversary. The second stage, *orientation*, is often considered the most important it consists in using cultural context to understand the worldview of the adversary. This worldview will become more and more accurate in subsequent reiteration and will help to decision-maker to take the right action. The third stage, *decide*, consists in deciding the course of action to pursue. The fourth stage, *act*, implies that after the decision is made, it is vital to act on it. The OODA loop helps to balance the need for making rapid decisions and the need for making informed decisions. This framework is relevant for this research because it gives a high-level description of the decision making the process a Blue Team member follows when investigating an attacker.

The NIST Cybersecurity Framework is a policy framework used to improve organizations ability to prevent, detect and respond to cyber-attacks. It organizes a list of activities into four categories: identify, protect, detect, respond and recover. Identify aims at identifying critical assets. Protect implements the mechanisms to ensure protection of the system. Detect implements the mechanisms and processes to spot cybersecurity events. Respond defines the activities to take action regarding a detected cybersecurity event. Recover develops and implements the activities needed to restore an organization's services after a cybersecurity incident.

The NICE framework is a NIST publication that describes and categorize cybersecurity work. It provides a common lexicon and taxonomy of knowledge, skills and capabilities needed to operate in the cybersecurity field. A typical use of this framework is to help the employers to profile and assess the workforce they need. It provides a common language that defines the work requirements for the professionals. It consists of a set of categories of cybersecurity functions, each with a subset of speciality areas, and each speciality area groups work roles identifying a set of

knowledge, skills and abilities required to perform the work role.

ISO 27001 is a standard that provides specifications for Information Security Management Systems (ISMS). It defines a six-step process that helps to define the scope of the ISMS and chose security controls to be implemented. The steps are: define a security policy, define the scope of the ISMS, conduct a risk assessment, manage identified risks, select controls to be implemented, and prepare a statement of applicability. With respect to the previously presented frameworks, ISO 27001 is less focused on the specific activities of the Blue Team, but more on managing the overall security posture of the organization. This, in turn, could be helpful in a later stage of this research to contextualize Blue Team's decisions.

NIST "Computer Security Incident Handling Guide" outlines the four-step process of the incident response lifecycle: Preparation, Detection and analysis, Containment eradication and recovery, and Post-incident activity. The preparation phase includes all the steps taken before the incident occurs. In the Detection phase, the events are analyzed to determine whether or not there is a security incident. During the third phase requires to interact with the system to contain further damage, then the root cause of the incident is investigated, and finally, the system is brought back to normal operational status. Finally, in the post-incident phase, the lessons learned are reviewed.

Summarizing, the OODA loop shows the decision making process of the Blue Team, and can help to understand how the Blue Team got to certain conclusions during an investigation. The incident response lifecycle describes the process followed by incident response teams. The NIST Cybersecurity Framework shows the main activities the Blue Team perform. The NICE frameworks precisely describe what the capabilities needed by a member of the Blue Team are. It can be used to support better profiling of the subjects of the research. Altogether the presented framework provides the background for further analyzing the Blue Team operations.

2.2.2 Elements of the SOC

SOC collects various suspicious alerts from sensors installed in the client network, then it correlates and analyze these events and eventually generates an alert for a security incident. Subsequently, a human analyst verifies the suspicious event and

decide if it is a true positive if it is the case the event is exposed to the decision-makers in a process called escalation [31]. As discussed at the beginning of this chapter SOCs fall into the class of Security monitoring, the class is at the intersection of the other three categories, and this is reflected in the variety of different functions carried out by a Security Operation Center. Those functions include: Log collection, Log retention and archival, Log analysis, monitoring security environments, Incident management, threat identification, and reporting. SOC is often described as a triad of elements that cooperate together: people, process and technology [32]. Therefore, the inner workings of the SOC will now be presented following this model.

People People working at SOC are divided between analysts and engineers. On the frontline of SOC, there is the Tier1 analyst. **Tier1** is a professional whose main duties in the SOC are to monitor the SIEM alerts, prioritize alerts, perform triage and decide whether or not a real security incident is happening. Then there are the **Tier2** analysts. Tier2 analysts are typically more experienced than Tier1s, and have knowledge in incident response, forensics and malware assessment. Their duties consist of receiving incidents from Tier1s and performing a deep analysis, identifying threat actors by correlating incidents with threat intelligence. They also decide how to proceed for containing and remediating a security incident. The **Tier3** analysts, also known as Subject Matter Experts or Threat Hunters, are similar to Tier2 but with more experience and even more knowledge. They are experienced in penetration testing, malware reverse engineering, and are capable of identifying and responding to new threats. Some of their duties include vulnerability assessments, reviewing industry news and threat intelligence data. They can also actively hunt for threats that infiltrated into the network. **Security engineers** are hardware or software specialists that focus on designing the security aspect of information systems. They can operate within the SOC or support their operations as part of the DevOps team. Finally, there is the **SOC manager** or Tier4 analyst, just like the Tier3 is a highly competent and skilled specialist, but operates on a strategic level, hiring, managing resources and the team.

Technology A Security Operation Center requires many different tools to effectively protect the system they are monitoring. Generating and collecting logs, as well as correlate events and generating alarms are some of the main tasks. A. Michail [32] identified the main tools that are part of every SOC platform and each tools supports a different purpose: Intrusion Detection System, Intrusion Prevention systems, and the most important Security Information and Event Management

Systems.

SIEM Is a technology that allows real-time analysis of security events (e.g. network traffic and logs) generated by the sensors placed within the organization's boundaries. It can be divided in two main components: a Security Information Management (SIM) and a Security Event Management (SEM) system. Where the former deals with log management whereas the later with real-time monitoring and incident management [32]. *Intrusion Detection Systems* (IDS) is technology that monitors the network for anomalies and suspicious behaviours. It can be further divided in NIDS (Network Intrusion Detection Systems) if they inspect network traffic, and in HIDS (Host-based Intrusion Detection Systems) if the monitoring happens on the hosts (e.g. resources being access and logging malicious behavior). Intrusion Prevention Systems (IPS) is a technology similar to IDS in the sense that they both monitor a specific source of events. However, IPS are not passive components as they can directly act on the threat (e.g. dropping packets or resetting connections) [32].

Process SOC's process defines the interaction between people and technology, within and to outside the SOC. They can be divided into four categories. There is a lack in the literature of precise definition of all these processes. Therefore they will be introduced without great detail. There are Business processes, Technology processes, Operational processes, Analytical processes. Here is an overview of these processes that was given in "Security Operation Center - A Business Perspective" [32]:

- Business processes define and document the administrative components required to efficiently operate a SOC while guaranteeing that the operations are aligned to organizational goals.
- Technology processes ensure that the IT infrastructure performs at optimal levels at any given time. They also maintain the information and document the actions pertaining to system configuration management, system administration, technology integration
- Operational process document and define the actions that are performed on a SOC on a day to day basis.
- Analytical process determines how security issues are detected and remediated. They also include the actions taken in order to learn about and understand surfacing threats.

2.2.3 Literature Review on SOC

In terms of academic research, the SOC is often analyzed from a business perspective and in terms of the human process behind it. An overview of the relevant research papers on the topic Security Operation Centers will now be presented. More specifically, papers related to the topic ”security analysts investigation process”.

One of the goals of this research is to understand the investigative process of SOC analysts better. Similar work has been done by Khalili et al [31], they tackled the need of improving security analysts performance by developing a tool that is able to monitor, measure, simulate and give feedback about SOC analysts. They identified the challenges of SOC as lack of a model that describes SOC analysis workflow, lack of tools to measure SOC performance, and lack of a convenient method to transfer knowledge amongst analysts. The authors identified eight investigation types divided into two categories, security-related incidents and policy violation incidents. The investigation types are then used to classify different tasks and activities, and finally their relationship is shown in a UML diagram. The authors concluded, showing that the system they designed has improved SOC performance.

An important step during security analysts workflow (more specifically, Tier1's workflow) is data triage. Zhong, Chen, et al [33] aimed at automating this process by studying security analysts' operation traces. They captured the traces of analysts operations performing data triage, then they created a graph representing the logical and temporal relationship of the events, finally from they used the graphs to construct a state machine. Their work demonstrated the feasibility of extracting a model for security analysts data triage process.

Another study that tackled security analyst workflow was done by Champion, Michael A., et al [34].The goal of their research was to understand the processes used by cybersecurity defence analysts in their job. They focus of their research was to identify team dynamics and factor influencing the team's performance. They demonstrated that effectively communicating teams are more successful than the one lacking communication skills. This proves that the security analysts investigation process should be seen as a part of a teamwork, and not as an individual effort.

Operational workflows of the SOC analysts have also been addressed by Sundaramurthy, Sathya Chandran, et al [35]. They acknowledged the fact that gathering insight on the operational workflow of SOC analysts can be a challenging task. Therefore, they adopted an anthropological approach by inserting in the SOC

computer science students trained on anthropological methods. This allowed the researchers to see the operational environment from the point of view of analysts.

The question ”how analysts think during an investigation” has been addressed in a research done by Sanders et al [36]. The authors investigated the cognitive processes of security analysts during the investigation process and proposed a model that explains such processes. They observed that the analyst's day-to-day work is mostly intuition-based. Even though intuition is mostly regarded as unreliable, the author argued that it plays a major role, so they proposed a model based on convergent and divergent thinking called the ambiguity-driven convergence model. The model shows that analysts are likely to rely on intuition first. When their intuition leads them into a high stakes situation, the analysts' tendency toward lower ambiguity tolerance results in the use of convergent and divergent thought processes to advance the investigation.

2.3 OPSEC

Defining OpSec Operation Security (OPSEC) is a classic military term that has been ported to the cyber security realm. OPSEC is about identify potential critical information, analyzing how adversary might learn this critical information, and taking the countermeasures required to prevent the adversary to interpreting or piecing together such information in time to be useful. This way OPSEC protects critical information from adversary observation and collection.

The OPSEC methodology was developed during the Vietnam War when it was discovered that public available information was analyzed by the enemy obtaining advanced information about certain combat operations [37]. Operation security is defined as the process used to identify, control and protect unclassified information of sensitive activities or operations. Once such information is identified it is possible to mitigate the threat or to deny a potential adversary the ability to compromise said operation. [38] [39]. This process is quite generic and is largely applied to a number of different fields such as military, business and cyber, or whenever there is a critical piece of information that has to be kept secret from the opponent. Operation security can be considered the complementary of intelligence gathering. Intelligence gathering focus on collecting information from different sources about a particular entity, and then to fuse this data to build an up-to-date and and correct view of the current situation [40]. OPSEC highlights the fact that intelligence gathering can

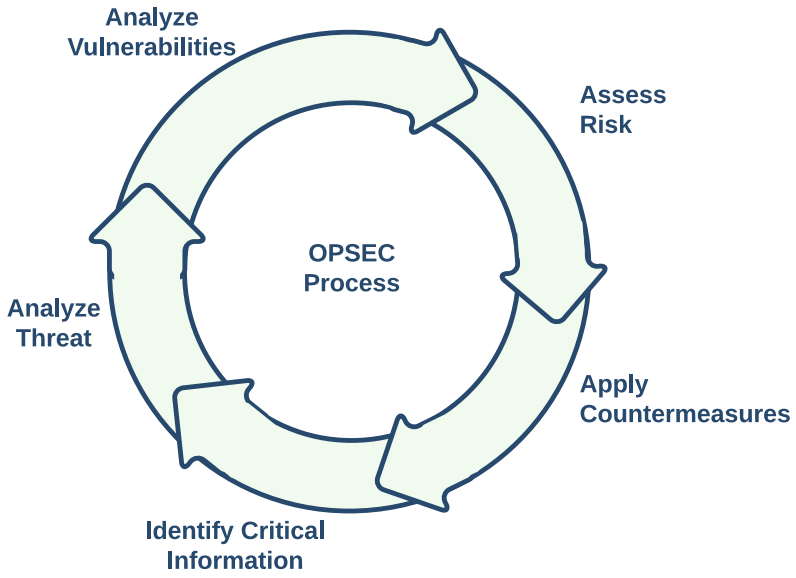


Figure 2.6: The 5-Step OPSEC process

be abused by an enemy, and publicly available information that is unclassified and apparently harmless can be aggregated forming the overall picture.

By definition, the OPSEC process involves five steps: identification of critical information, analysis of threats, analysis of vulnerabilities, assessment of risks, and application of appropriate countermeasures (see figure 2.6) [38] [39]. It is a general process that can be applied to any field in which an exist an adversary willing to gain the advantage, from the military to business and cyber. The first step identifies the critical information that if acquired by an adversary would cause arm to the organization. The second step implies understanding who are the adversaries so that it becomes clear what data they might be targeting. The third step helps to increase the visibility over an organization security exposure. The identified vulnerabilities are then evaluated in step four in order to understand the impact the exploitation of such vulnerabilities would have, and therefore being able to prioritize the efforts to mitigate them. The final step defines and implement countermeasures to protect against the threats. The OPSEC methodology is similar to Red Team operations in the sense that both try to figure out what an attacker could abuse certain information.

2.3.1 OPSEC Critiques

The idea behind OPSEC is to hide information to the enemy, in the cybersecurity realm this practice is referred to as security through obscurity. However, relying on the secrecy to achieve security is considered a bad security practice. The main criticism on OPSEC is that relying on information hiding goes against the Kerckhoffs's principle [8]. Also the NIST recommends against Security through obscurity "System security should not depend on the secrecy of the implementation or its components." [41]. The fallacy in this reasoning is that the Kerckoffs's principle is a cryptographic concept, and therefore it should not be applicable to operations. [42]. Asume the adversary know how the system works is different from conceal it. The efficacy of obscurity in operations security depends by whether the obscurity lives on top of other good security practices, or if it is being used alone.[8]

2.3.2 OPSEC Problem

Finding relevant research regarding Operation Security in the cyber realm is a challenging task. As was observed by the authors of "Cyber Deception Building the scientific foundation" the concept of "cyber operations security (OPSEC)" has had little systematic development or disciplined application in cyber security. The problem of Operations security is often tackled by giving a list of "best practices" to avoid disclosure of sensitive information, such practice is effective at an operational level by is insufficient to support a thorough study on OPSEC failures.

Another problem is that OPSEC may not always be desirable. For instance at a strategic level deterrence requires that the opponents have a clear insights into the intentions and capabilities of an organization [43], without such knowledge an adversary has no reason to avoid to attack. Theories from the discipline of "economics of cyber security" state that the optimal investment in security mechanisms is just high enough so that the cost of the attack is higher than the value of the "crown jewel" the enemy would acquire. The attacker therefore should be able to obtain enough information on the target to be convinced that penetrate their defences is not cost effective.

[8]The Kerckoffs's principle states that a system should be designed assuming that the enemy has complete knowledge of the system

2.3.3 Literature Review on OPSEC

As stated in section 2.3.2 academic research on OPSEC is really rare. After an extensive literature review only few papers were found addressing the topic of OPSEC, and none of them specifically addressing the OPSEC of security analysts. This demonstrate the importance of this research, as it may be the first step to fill this knowledge gap.

When it comes to OPSEC topics the focus of academic research seems to be mainly pointed at the issue of cyberattacks attribution. Researchers try to understand how to exploit adversary OPSEC failures or to identify attack patterns in order to attribute specific attacks to the appropriate threat actor. Wheeler and Larsen, 2003 [44] presented various techniques that can be used to determine an attacker identity and location based on traces that leaves on the system. Hunker et al, 2008 [45] suggested a methodology to identify attacker location based on IP. Rid and Buchanan, 2014 [46] discuss how to identify the country or organization behind an attack. Clark and Landau, 2011 [47] discuss how to trace cyberattack arguing that is more effective to investigate traces of the person performing the attack rather that traces of the machine.

Publication specific to OPSEC are often limited to the military field. A guideline published in 2007 define some OPSEC best practices in the cyber security field [48], it specifically address how to crate cyber OPSEC plan for control systems. An attempt on highlight the risk of OPSEC in the cyber domain has been made by Dressler, Judson et [49]. The authors demonstrated how it could be possible to retrieve information on sensible high level U.S. military members, they collected open available data from social media and used machine learning algorithms to correlated and extract valuable information.

Even if currently there is no research on the topic of OPSEC of security analysts, there is a considerable interest from the Red Team community on detecting Blue Team activities. A popular project that is focused on detecting traces of Blue Team investigation is Red Elk [50]. Red Elk is a SIEM [9] for Red Teams which is used to support Red Team operations by tracking Blue Team investigation and generating alarms. The tool collects specifics IOCs generated by the Blue Team, such as connections to Red Team servers or samples uploaded to public sandboxes, then it alerts the Red Team which in turn can make an informed decision on the next step

[9] A Security Information and Event Management (SIEM) is a software solution that aggregates and analyzes activity from different resources across the IT infrastructure

to take. The popularity of this project amongst ethical hackers shows that there is a push from the cybersecurity community to better understand how the security analysts can compromise their operational security and how attacker can possibly be able to detect it. However, both on the offensive and the defensive side there is still a lack of understanding of all the possible traces an analyst leaves behind.

Conclusion Maintain the secrecy of cyber investigation is a crucial part of cyber operations. The OPSEC process has been developed to help security operators avoid the disclosure of critical information to the adversary. This chapter discussed some critiques to OPSEC methodology and explained the importance of using OPSEC in cyber operations.

Another result of this section has been identifying a branch of research which is akin to identify Blue Team investigation. Academic research on attack attribution can be considered a different flavour of this topic in the sense that both aim to identify specific traces that can be attributed to one particular actor.

Despite the importance of maintaining the secrecy of cyber investigation, still little is known on the matter. An essential result of this preliminary research has been highlighting that there is a knowledge gap in the literature. There is a lack of a systematic analysis of the possible mistakes security analysts might do during their investigation. Moreover, it is still not clear to what extend security analysts are aware of the footprints they leave behind when they are investigating security events. For this reason, an additional research goal is to classify the possible indicators that are generated during an investigation.

2.4 Red and Blue Team interplay

In the previous chapters, it has been defined **who** are the main players in this research (the Red and the Blue Team), and **what** is the subject this research is analyzing (OPSEC). The next step is to determine **how** the two players interact with each other. For this purpose, the following section will offer an overview of relevant academic research that studied how Red and Blue Teams interplay.

The most comprehensive work that examines how the Blue and Red Team interact has been done by Shouhuai Xu in ”Cybersecurity Dynamics: A Foundation

for the Science of Cybersecurity " [51]. The author argues that modelling cybersecurity is more effective using a holistic approach rather than tackling the single building-blocks. The research is focused on modelling attack-defence interactions in cyberspace. The author proposes a set of metrics that describe the cybersecurity state, and explains how to identify the laws that govern the evolution of the cybersecurity state. Such laws are functions of cybersecurity metrics and time. The metrics proposed by the author are of five categories: metrics describing networks and configurations, describing human vulnerabilities, describing the defence employed, describing cyber-attacks, and for describing global security and situational awareness. These metrics are used as laws parameters to derive macroscopic phenomena form the underlying microscopic attack-defence interactions. This study calls to action other researchers into exploring the dynamics between attacker and defenders in the cyber domain, justifying the academic need for further research.

In another study He, Fei et al [52] applied game-theory to study the interdependency between service providers, attackers and defenders. The author designed a simultaneous game between the parties taking into consideration both defence strategies and attack strategies. The author studied different network topologies and evaluated how the success rate of defenders change based on differences in topologies and level of interdependency between elements of the network. The author demonstrates how some network topology were able to reach a Nash-Equilibrium between the attacker and defender. However, the model is still not mature enough to explain more complex configurations. This paper demonstrates how it is possible to approach the research on the Red and Blue Team utilizing mathematical models, which in turn can lead to a more systematic study of the topic.

Game theory has been used by Luh, Robert et al. [53] to derive a gamified model that defines attacker and defender interplay. The authors state that "the complex interplay of attack techniques and possible countermeasures makes it difficult to appropriately plan, implement, and evaluate an organization's "defence, for this reason, the model they proposed is based on a mapping of CAPEC[10] attack patterns to NIST SP800-53 controls[11]. They obtained a gamified meta-model that can be used to train personnel, assess risk mitigation strategies, and compute new attacker/defender scenarios in abstracted (IT) infrastructure. This study is relevant for two reasons. First, by mapping attack vectors to security controls, it lays the foundation for a comprehensive framework that incorporates both cyber offensive and

[10]Common Attack Pattern Enumeration and Classification (CAPEC) is a dictionary of known patterns of attack employed by adversaries. It is maintained by MITRE

[11]NIST Special Publication 800-53 provides a catalogue of security and privacy controls

cyber defensive measures. Second, the detailed game model proposed to capture the process of an attack-defence event, and this process has always been studied as two separate processes in the past.

Research has also tried to improve the integration of the two teams. D'Amico, Anita D. and K. Whitley [54] studied the methods, tools, and challenges of the Red and Blue Team in the U.S. Department of Defence (DoD) during integrated operations. The authors tried to understand how the two usually opposed teams can work together during missions that require integrated operations, such as incident response. After several interviews with the team members, the authors conclude that although there are cases of successful cooperation, there still need for improvement. Specifically, in data fusion, enhanced changed detection, incorporation of network maps, and better access tracking. This research highlights the need for further research on how to improve the integration between the Red and Blue Teams.

Talking about integration of the Red and Blue Team, Mattieau Branlat [55] studied the challenges faced by network defenders using a human-centred approach rather than a technological approach. The author examined a live event, with Red and Blue Team operating, and described the interrelated attack defence process. This study tackles, at the same time, the decision-making process of the attacker, and the decision-making process of the defender. The author identified core characteristics in the domain of cybersecurity that impacts the process of attack and defence, as well as their investigation. Those characteristics are uncertainty and complexity, joint activity (conflicts in the team), and adversarial nature. The importance of this research is due to its approach human-centred, which demonstrates that when dealing with investigation processes understanding the decision-making process is crucial. It is also important because it emphasized how cybersecurity is a fundamentally collaborative environment.

The importance of observing the Red and Blue Teams in a live environment has been remarked by many authors. Vykopal, Jan, et al [56] observed participants in cyber range and elaborated the lifecycle of cyber ranges. The lifecycle is made of five steps: preparation, dry run, execution, evaluation, and repetition. VISKY, Maj Gabor. [57] discussed the technicalities of developing a cyber-physical battlefield. The authors [58] are investigating new methodology for cyber exercises providing a framework for all aspects of an organization together and test their responses.

Conclusion The literature is very rich when it comes to study cyber defence exercises, probably because they provide an easier observation environment, and the results are easily reproduced. On the other hand, efforts to explain the relationship between the Red and Blue Team are much rarer. A new discipline,"Cybersecurity dynamics", that studies this relationship has been created , and it calls to action researchers to contribute to its development. This topic is backed up by some mathematical models based on game theory, however, due to the complexity of the topic and to the many factors that influence it those models are still not mature enough to be applied in the wild. The interrelationship has also been studied on a practical level by mapping attack vectors to security control, and so provide a good starting point for further research. Finally, the presented literature suggests that the study of the investigation process of the Blue Team requires a human-centred approach.

2.5 Ethical Issues

An important part of research in the field of computer science is ethics. Computer ethics is the science that studies how computing professionals should make decisions regarding professional and social conduct [59].

Even though ethics does not directly contribute to the results of the research, it is important to promote moral and social values in society with academic research. Academic research should always aim to improve society as a whole. However, trying to improve the efficiency of attackers by highlighting analysts' OPSEC failures might not be seen as having a positive impact on society. One question would arise: What are the ethical implications of improving attackers position? This section does not aim at answering this question right away, instead, it will provide some background research on how academics are questioning themselves on the issues created by studying offensive technologies.

The main topic of debate amongst academics is about teaching offensive technologies to students. Pike, Ronald E. [60] examined the issue of schools starting teaching ethical hacking in academic programs. The author argued that, even though ethical principles are thought to the students, they still lack the experience to apply those principles in practice. This study showed the importance of social circles and the value of professional networks.

Radziwill, Nicole, et al. [61] question if ethical hacking should be taught in the first

place. The author concluded that the industry pulls for more cybersecurity professionals, and the increasing potential risk determined general consensus that ethical hacking should be thought. The author also argues that regardless of hacking being taught or not hackers will still exist. Even if students are not taught ethics alongside hacking, they might as well learn that skills somewhere else without the ethical component.

In another study Hartley, Regina D. [62], defines ethical hacking pedagogy, moreover, he suggests ethical hacking as a computer security instruction methodology, and illustrates the ethical and legal consequences of teaching students to hack. The author sustains that security professionals need to have the same skill sets as attackers in order to recognize and defend networks from intrusion adequately.

Another field of research on the offensive side of computer ethics is vulnerability research. The author discusses the legitimacy of ”vulnerability researchers”. The author concludes that vulnerability research should be considered neither unethical or illegal. The main argument is that researchers are explicitly to prevent or mitigate harm occurring to third parties. Moreover, the author states that the researchers have a moral obligation to use their skill for the benefits of society. One argument against vulnerability research is that since it is against the law, it should be considered immoral. The author concludes by suggesting to craft several norms and guidelines to regulate vulnerability research.

Computer ethics is a very prolific field of research. Academics from all over the world question themselves trying to understand if certain research should be done, and if so, how to do it in an ethical manner. Teaching ethical hacking and research vulnerabilities have been largely discussed, nevertheless, the debate is not yet over. However, on the offensive side of computer ethics, we could not identify existing research that explicitly discusses the issues of a research that might benefit attackers. The lack of findings might be due to the need to explore the literature even further, or simply because no research has addressed this issue yet.

Chapter 3

Methodology

Section 1.2 explained what the research questions that will be answered in this research are. It also discussed their relevance and showed how they are related to each other. This chapter will explain the way each research questions will be answered. A mixed research method will be used; different research strategies and data collection methods will be adopted at each stage.

The research questions are following summarized:

- RQ1. How the Red Team can detect Blue Team OPSEC failures?
- RQ2. What are the most common SOC analysts OPSEC failures?
- RQ3. To what extent the Red Team is aware of the Blue Team investigation?

3.1 Problem identification and motivation

Problem Explication Identify OPSEC failures and measure their impact on Red Team actions is the problem that we will try to solve. According to the discipline of design science [63], a practical problem is a gap between the current state and a desirable state, as perceived by the participants in the practice. The desirable state is seen as better than the current one because it allows people to be more successful when engaging in the practice. The ability to spot Blue Team investigations through OPSEC failures can help the Red Team in many ways, for instance, it might help the Red Team to obtain access to the system, or to maintain access for longer, it might

also help them to create more insightful reports and hence improve the learning experience for the Blue Team. Likewise, also, the Blue Team will benefit from better understanding their common mistakes and how such mistakes can give an edge to the adversary.

Scientific contribution A design science project gets its basis from the scientific body of knowledge of the discipline and then produces scientific contributions that improve the body of knowledge itself. The result of the project is an artefact which is developed to address a practical problem. According to the authors of ”An introduction to design science” [63] there are different kinds of contributions that design science research can bring, which are: improvement, invention, routing design, and exaptation. The expected outcome of this research falls into the last category. Exaptation consists of adapting an existing solution to a new problem. The problem of detecting Blue Team activities can be considered a new problem in the sense that it was never perceived until a recent development of projects such as RedElk [50] which made the cybersecurity community realise that their practice could benefit from having these capabilities.

3.2 Research Strategy

The overall plan for this research will be explained in the following section. Many different empirical research strategies can be used, however considering the nature of the problem under analysis a mix of two different research strategies will be used, i.e. grounded theory and simulation. The need for using two different research strategies comes from the lack of previous research addressing this problem. It is necessary first to create the basic tools (the theories) and then observe how this theory applies. The former is reached with grounded theory and the latter with simulation.

The research will later present and discuss such theories and tools in details. For future reference, we can summarise the whole research as follow. Firstly OPSEC failures will be identified, then the assumption that these activities have an impact on the strategies of the Red Team will be proved (or disproved) during a simulated Red Team assessment.

Grounded theory is a research strategy that starts from the analysis of empiri-

cal data and proceeds bottom-up in order to generate theories. Grounded theory is particularly useful for exploratory research studies when a novel area of interest is addressed that requires the development of new theories. For this reason, grounded theory is well suited for this research [63]. Typically this kind of research starts by collecting data on a few objects and then based on the outcome; new objects are selected and investigated as long as they provide new insights. This iterative process is called theoretical sampling, and it continues until reaching the theoretical saturation when the data collected does not help to improve the theory further. Multiple analysts will be interviewed, and the questions answered to the next analyst will be refined based on the answers received from the previous one. This strategy will be used to guide the identification of possible flaws in the investigative process of SOC analysts. More precisely, it will be used to identify common OPSEC failures.

The second design strategy that will be used is *Simulation*. Simulations is a non-empirical research strategy that focuses on studying an imitation of real-world processes over time. The reason for using this strategy is that the artefact/theory generated using grounded theory will have to be validated. However, the nature of this research topic does not allow to validate the theories in the wild because it would present a series of challenges too difficult to overcome with the resources available for this study. One of such challenges is that observing the activities of the Blue Team implies alerting them that a Red Team assessment is coming, and as a matter of fact reducing the chances of success for the RT. The solution is, therefore, the use of role-play simulation or a war game between RT and BT. A role-play simulation is "a simulation, in which human participants take on different roles or profiles in the enactment of a process in a contrived setting" [63]. In summary, the simulation strategy will be used to satisfy steps five and six of the design science cycle, demonstration and evaluation of the artefact.

3.3 Data Collection Method

A critical aspect of this research study is the collection of the data. Being exploratory research, the kind of data needed is qualitative. Two types of data will be collected: the answers of the SOC analysts, and the actions of the Red Team in response to SOC analyst's investigations. The data will be collected in two main steps, which are: Identification of investigative process failures, Simulation & evaluation. The output of the former will be used to feed the second one. The output data of the two phases is a list of OPSEC failures and the impact that such failures have on the

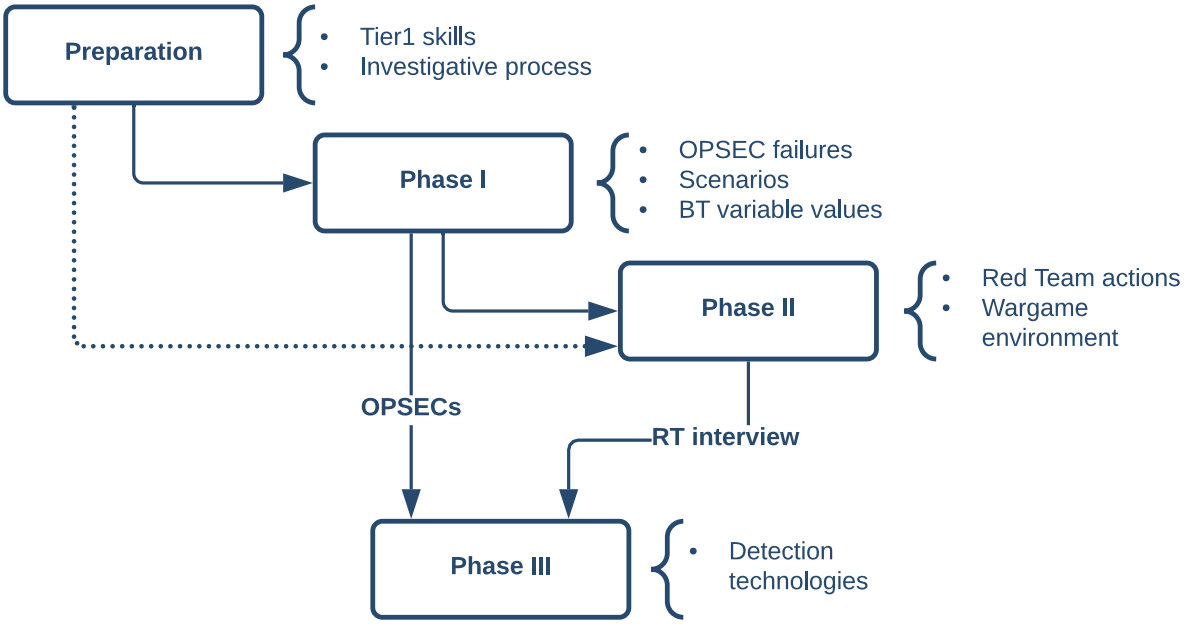


Figure 3.1: Mixed research Strategy

strategies of the Red Team.

In order to answer *RQ1* data collected in Phase I and Phase II will be combined. The OPSEC failures identified in Phase I and interview with the Red Team after Phase II will be used to evaluate there are technologies the Red Team can be used to detect Blue Team operations.

Preparation Formulating adequate questions for the interviews and designing scenarios for the subsequent data collection phases requires some background knowledge on the analyst's ways of work. For this purpose, the researcher worked as a SOC analyst for few months, observing and getting accustomed to the modality of operation in the Security Operations Center at Northwave [64]. This method was previously validated by Sundaramurthy et al. in "A Tale of Three Security Operation Centers" [35]. In that study a group of the authors' students with a blended anthropological and Computer Science background were hired in a Security Operation Center as security analysts, the researcher then analysed their on-field observations and interviewed them. Using this strategy, the researchers managed to have easier access to part of security analysts knowledge that would not be accessed otherwise.

In this case, the researcher acquired the knowledge necessary to work as a Tier1 analyst. However, it is worth mentioning that it is not in the scope of this research to analyse this knowledge as there are other studies which already collected data and

defined the workflow of a SOC analyst [31] [33] [34]. The goal of the preparation phase is to ease the work in the subsequent phases.

As Crandall et al. described in ”Working minds: A practitioner’s guide to cognitive task analysis” one of the challenges to perform a proper interview is that requires ”Well-trained interviewers with knowledge and skill that goes well beyond the understanding of standard data collection and analysis procedures” [65]. The preparation phase aims to provide the interviewer with such skills. Moreover, the author observed that ”Observation can be particularly effective when the researchers are well trained in the phenomenon they are studying and do not require a lot of structure for their data-collection activities” [65].

Once the preparation step is completed the knowledge acquired through the observations is used to formulate questions for the semi-structured interviews, and to design hypothetical scenarios. During the **Phase I (Investigative process definition)** step, the researcher will try to extract the embodied knowledge of the analysts.This is done by means of semi-structured interviews, and by proposing to the analysts a scenario and asking them to describe the actions performed in that hypothetical situation. Interviews are suitable for collecting data from people having advanced knowledge in a particular field [63]. Semi-structured interviews will be used because they are more flexible and allow the interviewee to answer more freely. Section4.2 will provide a detailed description of both the semi-structured interviews and the proposed scenarios, as well as the challenges faced when collecting the data. The results will then be transcribed for further analysis. According to grounded theory research strategy, the qualitative data collected is then analysed through coding and categorisation. The researcher will identify pieces of data from the scripted interview and assign a label to each one. The coded data will then analysed and used to identify a list of possible IOC (Indicators of compromised) that delineate the footprint of a security analyst.

The final step in data collection is **Phase II (Simulation and Evaluation)**, it will be used to observe the interaction of the two teams in a realistic situation. As discussed in previous chapters one of the limitations in studying the actions of the Red Team is how to observe them in their ”natural environment”. A wargame will be used to overcome this limitation of observing the attacker and defender interaction. In such simulations Red Teams will try to break into a system protected by a Blue Team while also using the detection tool to try detect their activities. For this purpose a virtual environment will be designed and created where the two teams can work. This stage requires the definition of the requirements, identification of appropriate

technology to use, and identification of the elements to include in a realistic environment. An additional goal is to demonstrate that it is possible to perform this kind of research in a short amount of time and with limited resources.

In **Phase III** we will reason on the findings of the previous phases to reach new conclusions. For this reason, there is no data collection method associated with this phase

3.4 Preparation

The first issue encountered in phase I trying to identify OPSEC failures through semi-structured interviews was the need to create a common ground upon which build the conversation. However, a language barrier exists between highly specialised workers and common people. The vast majority of specialists are used to discuss technical matters with other specialists, and for this purpose, they often use technical terms, or they refer to situations and tools specific for their daily job. Such technical lingo is called technical jargon, and it has been confirmed to be a communication barrier in many researches [66].

For this reason, having a technical conversation with a specialist can be a challenging task if the interviewer is not a specialist himself. The task becomes even more challenging when considering semi-structured interviews. In this kind of semi-structured interviews it is necessary to completely understand what the interviewee is saying in order to be able to maintain the conversation on the right track and be able to pivot to a different topic if needed. Moreover, holding a conversation with another specialist removes the need to translate “on the fly” technical terms or to explain the meaning of the unknown, resulting in a conversation which is more fluid and rich in information. As remarked before a similar technique was also used by Sundaramurthy et al [35], which introduced a group of students to work as SOC analysts while they also made anthropological observations on their job and the other analysts. Using this innovative technique the author managed to obtain more in-depth insights than with observations done as “outsiders”.

In the case of this research having direct experience of the phenomenon helped to approach the topic with a fresh perspective and generate more relevant theories and hypotheses. This experience contributed to the research as follows:

- Being trained as SOC analysts reduced the uncertainty on which kind of answer the other analysts might give.

- Helped at identifying questions that would provide precisely the information needed.

- Allowed the other analysts to talk more freely without them concerning about explaining trivial concepts.

- Allowed to perform a preliminary observation on the work of SOC analysts

- Helped acquire the necessary technical skills to perform an investigation.

Finally, being trained as Tier1 was beneficial for Phase II. We needed precise control over the action of the Blue Team and the ability to perform a full SOC investigation.

3.5 Phase I

The goal of the first phase is to extract the embodied knowledge of security analysts. “Embodied knowledge, sometimes called tacit knowledge, is situated in the minds of people and is often difficult to formulate in an explicit way” [63]. Using semi-structured interviews security analysts have been interviewed to elicit their knowledge. They have been selected using criterion-based sampling. This collection method is widely used for the identification and selection of candidates for the most effective use of limited resources. This involves identifying and selecting individuals or groups of individuals that are especially knowledgeable about or experienced with a phenomenon of interest [67]. Because a single researcher has conducted this research and in the limited period of six months, it was not possible to interview a high number of security analysts. Therefore we selected the candidates based on their experience as security analysts and on the fact that they were currently employed as analysts. The reason for this last requirement is that the work of the analyst is continuously evolving; therefore, it is important to access recent knowledge. The negative aspect of this sampling method is that it is less generalizable due to potential biases that can be introduced in the selection process.

3.5.1 Interviews

The interviews have been conducted in two phases. The first phase was a simil-brainstorming session to stimulate the creativity of the interviewee and identify as many OPSEC failures as possible. As described highlighted in Phase I analysts are extremely fast at processing the alarms; however, identification of possible mistakes requires them to slowly think about the way they perform their analysis. For this reason, the first set of questions served a double purpose. The first purpose was to profile the interviewee and get some general information about its experience and the way he performed the analysis. However, the questions were organized in a way to start from a high-level description of their process to a detailed analysis of the steps they follow, in such a way serving the second purpose of slowing down their cognitive process and helping them in better understanding and explaining the reason behind each step they made. Understanding the reasons why they made certain choices eases the process of identifying possible mistakes. The result of this process was that certain choices and heuristics the analyst used almost in an unconscious way are now clear in their mind. In appendix A, the set of questions are listed.

3.5.2 Hypothetical scenarios

The second part of the interview is the hypothetical scenario study. The scenarios were designed using the knowledge acquired from the Phase I and in collaboration with the Red Team of Northwave, which highlighted the most common tactics and techniques they use during an assessment.

At this stage of the interview, the analyst has already answered the previous set of questions and has fresh in his mind the unique process he follows. So they are presented with a set of hypothetical scenarios and asked to walk through their analysis describing what they were observing, what they were assuming and thinking about the event, how they were collecting additional information for context, how they were trying to maintain their OPSEC, they were also explaining which actions they would take and why. The output of this first step in the hypothetical scenario analysis contributes to defines a baseline of actions that can be performed during an investigation in a similar situation. Such set of actions will then be used the create the independent variable[1] for the controlled experiment in Phase II. Once

[1] In

the analysts completed the analysis of the alarm the hypothetical scenario analysis continued by asking to perform the very same analysis again but this time assuming to be an analyst who is totally unaware of OPSEC practices and therefore making every kind possible mistake. The analysts were then asked to describe what they could do wrong and the reasons why such action constitutes an OPSEC failure. The first analysis served the purpose of highlighting the measures the analysts use to preserve their OPSEC while the second analysis highlighted the mistakes they make. Thinking about others mistakes is a technique known as Elicitation by critiquing (EBC), and is a cognitive task analysis methodology that take advantage of the experts ability of analyze another's work. The use of knowledge elicitation methods such as critiquing technique to analyse the work of analysts has been proposed by Miller et al. [68].

3.6 Phase II

The data collection method used at this stage is a type of simulation. The goal is to observe the reactions of the Red Team to Blue Team activities. As was explained in a previous section, it is not feasible to achieve this goal in the real word without informing the Blue Team about the Red Team assessment going on and therefore prejudice the validity of assessment for the customer. For this reason, it is necessary to organize a cyber-wargame. In this way, it is also possible to reduce the uncertainty related to this kind of study by creating using two controlled variables. The actions of the Blue Team and of the regular users will be controlled by the researcher and will follow a scenario defined before the actual wargame.

The use of wargames between offensive and defensive teams in the cyber domain has been proven effective to test the effectiveness of new techniques for the Red Team. Heckman et. al [69] during a real-time, Red Team/Blue Team cyber-wargame experiment organized by MITRE. The Blue Team even used deception techniques to trick the Red Team into believing fake information was real demonstrating the advantages of having control over one of the teams. Attempting to study the interplay between attack and defence, also M. Branlat [55] adopted the strategy of using cyber wargaming to observe Red and Blue Team in a realistic scenario.

3.6.1 Wargame

We will organize a wargame with the Red Team of Northwave. In this simulation, the members of the Red Team will play as the attackers and will try different techniques to get access to the system. On the other side, the role of the Blue Team will be played by the researcher itself and some collaborators. The Blue Team and regular users activities are the controlled variable of this simulation. The role of the defence team is to both perform the investigation and to emulate actions of regular users so that the Red Team can try to gain access to the system. For example, if the Red Team tries to send a phishing email, the defence team will emulate the action of a regular user and click on links and provide credentials if required. In this case, the Blue Team will start sending signals to the Red Team in the form of OPSEC failures, such as trying to connect to their infrastructure with an identifiable IP address, and the responsiveness of the Red Team to such stimuli will be observed. The Red Team will be asked to keep notes of the actions they take during the exercise. Those notes will then be used to map their attack paths, and to compare their actions to the actions of the Blue Team in order to determine if the SOC investigation influences the Red Team.

3.6.2 Infrastructure

Organizing a wargame ,even if it is just for a small number of people, requires a solid infrastructure. The Red Team will obviously try to infiltrate the system with as many means as possible, for this reason it is important to have a high level of awareness of the possible vulnerabilities in the system. However, identifying all the possible holes in the security environment would require a lot of effort and skills beyond the possibilities of this research. One advantage of using an automated environment is that it is possible to easily recreate the same conditions in subsequent experiment, this in turns means that at each iteration it is possible to tune the system according to the findings of the Red Team. With this strategy, after a certain number of times the same environment has been used it is possible to obtain a system which is very secure, and in which only the intended attack paths are exploitable.

The are other researches that studied different aspects of cyber attack and defence [55] [70]. However, those researches were either performed by observing specific aspects of an already organized wargame [55] or used a complex infrastructure which required a lot of resources, maintenance and planning to make it work. In

both cases, it results in the difficulty to quickly set up an experiment to get further insights. The infrastructure that will be used in this experiment aims to overcome such difficulties.

3.7 Phase III

The conclusive phase of the research is Phase III. In this last phase, first, the infrastructure supporting Red Team operations will be briefly studied to identify what are the elements that the Blue Team might investigate. This first step is essential because, in order to be able to set up an effective monitoring system for the Red Team, it is necessary to understand what can be monitored. Subsequently, different technologies that can be used to detect OPSEC failures or that can be used to monitor the Red Team infrastructure will be presented.

3.8 Limitations

This section will discuss the limitation of the presented research strategy and data collection method. Following are summarized the main limitations:

- The members of only one SOC were interviewed therefore the findings are hard to generalize.
- Due to the participant observation in phase I the researcher worldview might have influenced the way data is evaluated
- The research is influenced by the personal belief of what the researcher thinks is important
- Identify all the requirements for a realistic environment might require an additional research in itself
- Insights on OPSEC failures are limited to the scenarios which have been analysed, but there could be more OPSEC failures related to different situations.
- A single iteration of the wargame environment was planned due to time limitations, and therefore many unintended attack paths might still be present

Chapter 4

Analyst OPSEC

The previous chapter described the main challenges in answering the research questions and gave an overview of how each phase has been planned to overcome the challenges and reach an answer. The first of the challenges is the identification of OPSEC failures. Even though the analyst workflow has been addressed before [71] [36] [31] [72] [73] [74], no previous research was found to perform an in-depth study on the mistakes analysts make during investigations. The following chapter will describe how by using a mixed approach of different interview techniques, the most common OPSEC failures have been identified. Moreover, it will also discuss how the additional information acquired during the interviews on the analyst investigative process has been combined with the OPSEC failures to define the Blue Team's action independent variable which is used during the wargame in phase II.

4.1 Building Blocks

This research is rooted in computer science; however, the answers that we seek can not be found by just looking at the technical side of cybersecurity but also at the human side should be analysed. As a result, technical and cognitive aspects are blended together. The following section will introduce the non-technical building blocks necessary to better interpret the results of the interviews.

4.1.1 Cognitive Task analysis

Cognitive Task Analysis (CTA) attempts to explain the mental processes involved in performing a task, and it is used for studying and describing reasoning and knowledge. The mental processes include the knowledge, skills and strategies that are needed to accomplish the task functions. In other words, Cognitive task analysis is the discipline that studies and captures the “how” and “why” the mind of the expert works [65].

CTA is used to describe the cognitive aspects required to properly perform certain tasks, and is not related to use of specific software or tools. Wei and Salvendy [75] stated that the goal of CTA is to “emphasize the knowledge base for the whole job: its organization and the interrelations among concepts or knowledge elements”. CTA is used for a variety of different purposes, however for the scope of this research it will be used to elicit the practitioner’s mental representation of knowledge.

There are a variety of different techniques available in the tool belt of a CTA researcher to reveal such hidden knowledge. “Observation and interviews” are often used in the initial phase of CTA when the domain is still young and needs to be explored. “Process tracing” this method is more formal than the previous one and explores the cognitive structure and processes underlying task performance, it is used when we can easily define a task that is representative of the actual task scenario, an example is the analysis of verbal reports and protocol analysis. Protocol analysis refers to having persons think aloud while performing or describing a task and then using verbalisation to infer subjects’ cognitive processing. The third set of techniques are called “conceptual techniques” and are an indirect method used to analyse a large amount of data to find interrelation to analysis tasks. This method is used when domain knowledge, structures, interrelations of tasks need to be defined and known; some examples are conceptual graph analysis, error analysis and questionnaires. Lastly, “Formal models” use models and simulations, and are used when a task needs quantitative predication [75].

It is worth noting that the methods of the first category require the observer to be trained in the domain of knowledge. In order to improve the effectiveness of CTA, it is often advised to combine at least two methods of the above mentioned. Techniques of the first and second family of methods are used in this research.

4.1.2 Knowledge Transfer

An important concept that is important to grasp is knowledge transfer. Knowledge transfer is the main subject of study in the discipline of knowledge management and was defined by Argote Ingram as “the process through which one unit (e.g., group, department, or division) is affected by the experience of another” [76]. This process is used to transfer knowledge from experts to novices. There are typically two kinds of knowledge that are of interests in the context of IT, i.e. Encoded and Embedded knowledge. Encoded knowledge is a type of knowledge where the information is expressed in terms of a common ‘language’ understood by both the expert and the novice [77]; it can be code, statics, books or manuals. Whereas Embedded knowledge is not explicit, it is usually locked in processes, products, culture, routines, artefacts, or structures [78].

As was observed by Swap et at [79], knowledge is often embodied in individuals, and there are some advantages in doing so. Experts can describe their knowledge and adapt this description to depending on the listeners, hence spreading it more effectively. On the other hand, in this process, they can not perform their daily work properly. Moreover, if they leave the company, such knowledge leaves with them. Those are important reasons to capture such knowledge, in particular when it comes within the context of analysts investigations.

The reason knowledge transfer has been introduced in this chapter is because one of the main reasons analysts make OPSEC mistakes is because they might not have experienced or learned about certain aspects of it. It is essential to understand how senior analysts acquired their knowledge about OPSEC best practices and how they could transfer it to newcomers. It is also a step towards finding a way to transform embedded knowledge (or tacit knowledge) is encoded knowledge so that it can be more easily shared in the community.

4.2 SOC Analysts Interviews

As it was discussed in section 3.5, a semi-structured interview has been carried out. Nine analysts have been interviewed for a total of 13h of recorded material. More details regarding the participants, their answers and the question asked can be found in appendix A. The interview resulted in many different findings, some of which confirm the results of previous researchers and others which are new. Among

the results that confirm existing research, there is a description of the investigative process of the analyst. Among the new result, there are insights on the training process of SOC analysts, and the description of possible mistakes an analyst can do (OPSEC failures). Each one of these results is now presented.

Analyst Training Process

One of the first questions asked to the analysts was *Which kind of SOC/analyst related training do you have?*. The goal of the question was to explore how the analysts are trained for their jobs, and how they pass down knowledge to other analysts. Most of the interviewee pointed out that the training they received consisted of a general introduction on the tools and the process, and subsequently, they ”looked over the shoulder” of another analyst. This process of transferring knowledge by looking at the actions of others is very common in many industries.

While this method is common and broadly used what a new analyst learns depends directly from the skills the senior analysts have. Some of the interviewees pointed out that this could be a possible problem during the training process. New analysts often might receive training from different seniors depending on their availability, each one of whom has developed a different investigative process over the years. This can be a cause of confusion for a new analyst. On the other hand, other analysts mentioned that this could benefit the new analysts as it is possible to experience the same task from different angles, and consequently come up with a more complete process by filling any gap with the knowledge from the other analyst.

Some of the more experienced analysts who were interviewed came from different backgrounds and had worked in different roles for different companies before ending up working as analysts. They said that their previous experience often helped them to reach a conclusion faster than they would otherwise. However, also the less experienced analysts benefit from the knowledge they had acquired outside the boundaries of their regular job, for instance, related to their hobby or interests in the IT field such as network analysis training, cracked games, CTF[1], ethical hacking experience.

The difference between what a novice can learn from different senior analysts is an issue of knowledge transfer. As was highlighted by Khalili M. [31] analysts usually

[1]Capture the Flag (CTF) are hacking competitions where the partecipants have to obtain ”flags” which are secrets hidden in purposefully-vulnerable programs or websites

possess different knowledge, since they gain different knowledge during each investigation related to different clients. Observing different analysts can help mitigate this issue. However, especially for learning about OPSEC practices, each analyst often needs to experience some mistakes first hand to build the skill-set necessary to decide what to do and what not to do. This confirms the basic assumption of this research, i.e. that there is not a convenient way to transfer this kind of knowledge between analysts.

Investigative Process

The previous question aimed at understanding the background of the analyst, along with discovering potentially unknown sources of information where the analyst could have acquired certain skills (perhaps OPSEC related). The subsequent question aimed to explore possible differences, if any exist, between the standard investigative process that was described in previous research and the actual process used by the analyst. The analysts were asked to provide a high-level description of their investigative process. They were then presented with the Blue Team kill chain proposed by Kumar et al. [74] and asked if the scheme matches what they do on a daily basis (see figure 4.1).

The use of a scheme or task diagram was proposed by Militello & Hutton [80] as one of the CTA strategies to identify tasks that have a high cognitive requirement that can then be further investigated. In their work, they proposed to present the subjects with a broad overview of the tasks and asked them to highlight the difficult cognitive portions in it. Such portions can be subsequently broken down in subtasks to refine the research better and obtain a better model.

All the interviewee provide a great deal of information to this question describing in detail both the steps they perform to investigate and additional tasks that might be required to support the steps. Interestingly the answers of the analysts were quite uniform, and no one mentioned any action they usually do, which was not already present in the scheme. Despite all information provided at this stage, it was decided not to analyze that further as it was not directly contributing to answering the research question, besides they just confirmed what had already been discovered in previous research.

After further inquiring with some of the interviewees they highlighted the fact that thanks to automation and the tools they use most of the security events go

Figure 4.1: Blue Team Cyber Kill Chain

through the first three steps (i.e. gather, detect, alert) and are filtered automatically. However, such steps might, in some edge cases, require human intervention. The steps where they spend most of their time are triage and context, whereas plan and execution are performed almost simultaneously.

The various descriptions of the investigative process have been combined and the result is the following reported: *"Look at the console and quickly scan through the queue to determine which alarm has higher priority and needs to be escalated first. The decision is largely based on the analyst's experience, and on the fact that some alarms might be related to some event they already investigated and of which they already know the outcome. This happens because sometimes, the same high-level event might generate many different alarms. Once they decide which alarm to investigate, they proceed to get situation awareness, trying to understand what the alarm is telling and on what rule is triggered. They try to see if it can be related to some other on-going investigation triggered by other previous alarms or to just look for similar alarms which might have triggered together. Next, they try to gather more context by using a number of different tools, but foremost by looking at existing logs. This step has been pointed out by some of the interviewees as where an OPSEC related mistake might actually happen, especially if logs are not sufficient and they need to dig additional information from other sources. Finally, once they decide if the alarm is an actual incident, they proceed with the escalation to the customer. In the case that it is not clear how to proceed or if they unsure of the analysis they always consult another analyst*

Interestingly the description of the investigative process provided by each of subjects, despite having slight differences between each other (especially in the moment when to consult another analyst), is accurately described by the generalized workflow diagram for analysts investigations proposed by d'Amico et al in "Achieving cyber defense situational awareness: a cognitive task analysis of information assurance analysts" [71] (See figure 4.2). This observation contributes to validating the findings of their previous research, mainly because despite the fact that different data collection methods were used it resulted in the same result.

Another issue that can be deduced from the interviews is that analysts create

site-specific knowledge overtime. Such knowledge is a mental model that defines what normal behaviour for a specific customer, including the type of alarms that trigger more often is, and unconsciously the analyst looks for deviation from the standing behaviour they are accustomed to . As a result, the analysts struggle to explain certain decisions (or what is considered normal) to the others, especially to novices with less experience. This issue was also highlighted by D'Amico et al [71]. The site-specific mental model is a heuristic that evolves over time and has the most impact in the initial phase of situational awareness when the analyst is prioritizing the alarms.

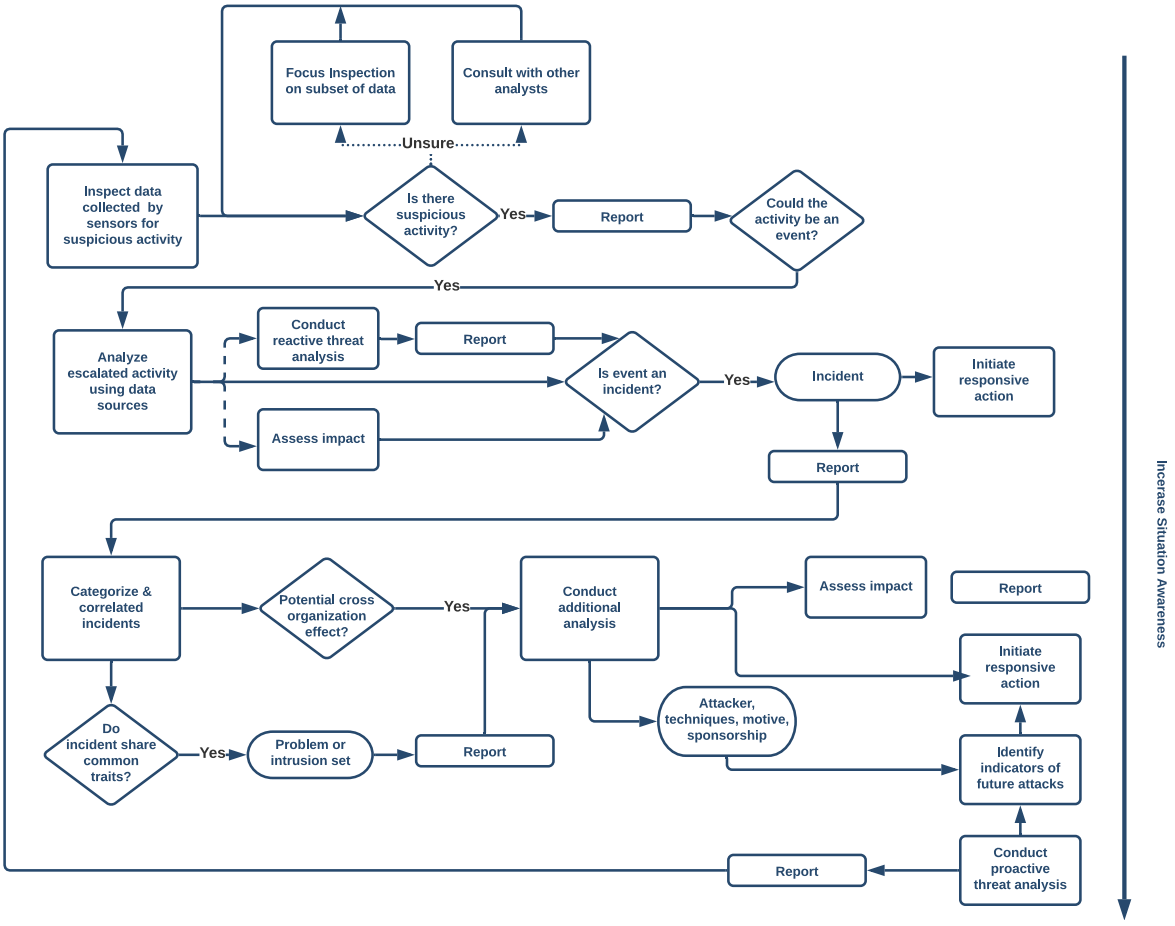


Figure 4.2: Analyst investigation generalized workflow

4.3 Hypothetical Scenarios Analysis

As it was explained in the section 3.5, the previous set of questions was designed to solve the problem that "Experts don't know, what they know". In the time they provided answers to previous questions, they had the time to slow down their process

and break it down in the single actions they perform. They also had the time to get in the mindset of thinking about the reasons why they perform certain tasks in certain ways. In this way, it was possible to obtain more insights from them and to ease the access to that part of their knowledge that is mostly unconscious.

The hypothetical scenario analysis is a knowledge elicitation technique which has been previously used by d'Amico et al. In their work they used this technique to overcome the issue of confidentiality and classification which prevented the subjects from sharing details of the cases they were working on [71]. The technique involves working with the analyst to analyze an imaginary situation involving a real cyber attack.

The knowledge elicitation technique is enriched by combining a "think aloud" method of CTA to obtain more insights. The result is a two steps hypothetical scenario analysis, where in the first step the analysts explain their actions in the hypothetical situation, and in the second step, they get to describe what a non-experienced (or novice) analyst would do in the same situation with emphasis on the mistakes that could be made.

4.3.1 Good OPSEC scenarios

The analysts were presented with five scenarios. Each scenario has been designed in order to cover a different part of a hypothetical attack path that attackers could follow, starting from outside trying to get the initial foothold up to the activities they would perform once they have established persistence on the system. The five scenarios are: Phishing email, fileless attack with word macro, command and control traffic, password spray attack, internal calls of a malware detected by HIDS[2]. More details about each analyzed scenario can be found in appendix A.

The subjects performed the analysis as it was in front of their SOC console, saying aloud the particular actions they take, the IoC[3] they are looking for, the assumptions they have when they start the investigation and how such assumptions can change based on hypothetical IoCs they found. Most importantly, they were asked to specify which (if any) measures they take to prevent their OPSEC from

[2]Host based intrusion detection systems are IDS capable of monitoring and analyzing the internals of a computing system as well as the network packets on its network interfaces

[3]Indicators of compromise (IOCs) are defined as of data that identify potentially malicious activity on a system or network."

being compromised.

The responses provided to the first three scenarios were in general, both breadth and deep. It was deep in the sense that the analysts explained in detail each IoC they were giving attention to and why they were considering it. For example, in the phishing emails, they looked for suspicious links, performed analysis on the landing webpage, looked for suspicious information in the header of the emails and evaluated the language used to write the email. Their responses were also breadth in the sense that they asked many if-then, therefore also the resulting description of their responses were rich in branches. On the other hand, the more the scenarios were set at advanced stages of the cyber kill chain the less “rich” the description of their actions were, and signs of suspicious activity directly led to an escalation without many steps in between. There are many reasons for that. The first one is the fact that the paradigm changes from an external attempt of infiltration to an assume-breach scenario where the tolerance to risk is much lower. The result observed is that they switch from a mindset of gathering context to take a decision to a mindset of gathering context to help eradicate/contain the problem. The lack of information provided in the scenario, the more at an advanced stage an attack is the more contextual information the analyst needs to reach a proper decision. Another reason is that the deeper in the system the attacker is, the more the tasks needed to handle the investigation are part of the skill-set of the threat hunter rather than of tier1 analyst. The last reason is the lack of exposition of the analyst to similar situations. As a matter of fact, it is more common to see multiple attempts of external attempts to infiltrate then to observe an attacker already inside, and the reason for that can be related to the presence of the SOC itself as nearly all the attacks are stopped in the early stages.

When the analysts were asked which OPSEC measures they take to prevent the compromise of their operations, the first reaction observed was a struggle to pinpoint which such measures are. This can be attributed partially to the fact that the interviewee question was not clear enough. However, a major role played the fact that such measures are often used as part of routines and good practices that “just make sense”. This initial struggle is a confirmation that OPSEC measures are actually part of the professionals’ knowledge which is unconscious, or as part of unconscious competence.

Subsequently, for each scenario analysts provided a list of things they do before, during and after the investigation to avoid leak information to the opponent. The first is to always use a VPN before starting anything, this way hiding their IP the attacker

is not alerted about their presence as the SOC IP might be known. Whenever they are required to use public services to dig more information or scan a website, they perform a private scan which hides the results from the public. Otherwise, an attacker who is monitoring these services can identify that the phishing landing page has been detected. Another OPSEC measure is always to avoid uploading malware samples to scan engines, as the results are public and can be easily queried by an attacker. If a malware requires further investigation, it is executed in a sandbox environment; however, this technique requires special precautions. A malware might need network connectivity to function properly and reveal it is malicious intent, and this might reveal the IP address of the analyst. Another option is that the malware might use special methods to detect if being run in a sandbox and alert the attacker.

It can be observed that all the answers given are specific to certain situations, such as uploading of samples and establishing connections to suspicious IP addresses. Not all the interviewee proposed the same OPSEC measures. The analysts who had some degree of experience as threat hunters pointed out additional measures, although also these were very specific to certain scenarios. Some proposed that it should be avoided to share links in message applications used to communicate between analysts as those have a function of preloading the image to display some of the content. In the last scenario, it was proposed to avoid doing changes to the system to perform the investigation (such as creating new accounts to perform certain tasks) because if the attacker has already established a presence in the system, then all these actions can be detected. However, in the latter case, this kind of changes to the system are outside the tasks a tier1 is normally allowed to do. When the last analysts interviewed were asked about the OPSEC measures proposed by the one with threat hunting experience, they confirmed that it was something they might have come up with given more time. This confirms that their answers are dependent on each individual scenario analyzed. It is a safe assumption to say that the use of different scenarios could have revealed additional measures other than the ones already mentioned. Therefore, it is left to future research to come up with alternative scenarios to identify additional OPSEC “good practices”.

4.3.2 Bad OPSEC scenarios

In the second part of the hypothetical scenario analysis, the interviewee was asked to ”put himself in the shoes” of untrained SOC analyst. They were required to perform the whole investigation again as-if they were not skilled analysts, and empha-

sise which OPSEC related mistake they would make.

Each OPSEC mistake can have a different impact on the operations and affect different aspects of it. For this reason, three categories of OPSEC failures have been identified based on the aspect of the (investigation):

1. OPSEC failures which can **compromise the investigation**
2. OPSEC failures which can be a **risk for the analyst**
3. OPSEC failures which can be are a **indirect risk for the customer**

This result has been achieved by analyzing the interviews and labelling the pieces of information the interviewee said. Moreover, it has combined the "bad" OPSEC behaviour with the "good practices" mentioned in the first half of the hypothetical scenario analysis to obtain a list of OPSEC failures. As it was previously mentioned such a list is specific to the scenarios which have been analyzed together with the analysts, and therefore perform the same analysis again but using different scenarios can result in a more exhaustive list.

The first categories include all the OPSEC failures which can give away the status of the Blue Team's operation to the adversary. The second category includes actions which directly exposed the SOC system to external attacks. Finally, the third one includes the actions which can compromise critical business functions of the customers as a consequence of an analyst's actions. Note that a logic category which has not been included is OPSEC failures which can be a direct risk for the customer, the reason for that is it would not have any value to this segmentation as all the OPSEC failures translate into a direct risk for the customer. Following the elements of each category are listed and explained.

Compromise the investigation

Execute a malware in a non sandboxed environment. A sandbox system provides a safe environment where users interact with the malware without consequences, however (although very unlikely) it is possible, for instance that an inattentive analyst opens an email attachment which contains malicious Microsoft macros directly on the machine he is working on.

Establish a **connection to a suspicious IP without VPN**. A VPN can hide the real IP addresses of the analysts. Highly motivated attackers could monitor their own infrastructure looking for suspicious connections that do not match their expectations.

Provide **clearly fake credentials** (it's important to remember that the adversary is not blindly doing campaigns). When the analyst is trying to obtain more context for the analysis, he might need to interact with the phishing webpage to verify if there is actual malicious intent. However, in doing so it might be necessary to input credentials to login pages. Using clear fake credentials or standard keywords that can be traced to the SOC corporate identity might alert an attacker who is looking at what the victims input.

Uploading of **malware samples to online services**. Services such virustotal allow users to upload samples of malware that are then tested against multiple AVs and the results are made available to the community. However, such services also provide simple queries which can be used to retrieve information from these services. An attacker could query for the hash of his malware and when a result is returned then someone must have uploaded the sample.

Use **public scans for websites**. Other services scan and analyse unknown URLs in order to identify potentially malicious websites. However, the results of the scan are available to the public by default and can be queried by an attacker who want to see if his phishing website has already been detected.

Creation of **artefacts on a compromised system**. Even though this highly depends on the kind of SOC and agreement with the customer, there are cases when the borderline between threat hunters and analysts blurs. In these cases, the analysts might decide to perform some changes on the system they are investigating to gather more context. For example, the creation of a bogus user with elevated privileges could be created for the SOC so that the analysts can access parts of the system which are normally restricted and perform a more thorough analysis. However, an attacker who has already gained access to the system can monitor such changes and be alerted by such suspicious activity.

Sharing of links via chat messages apps. Some popular chat messages preload the content of the url in order to show a preview of the web page whenever a link is shared. As described in section 4.2, communication among analysts is very important, and it is often the case that they share suspicious links via chat message seeking advice from other analysts. However, the receiving party might not have

been connected to the VPN when the message is received, hence a connection is created in the background which can in turn alert the attacker.

Communication with the customer **through compromised systems**. The time between when an attack is performed and when it is detected is often very short, however it is important to always switch to an assumed breach scenario mindset when something happens. It might be the case that when the analyst escalates the incident to the customer their email is already compromised, or the system from which the system is opened is already compromised. In this case the analyst is not only alerting the customer but also the attacker itself of the result of their analysis together with all the details of what has been observed and when.

Risk for the analyst

Provide own credentials instead of fake credentials. Especially when trying to quickly get through an analysis a distracted analyst might not have a fake account ready for access a certain website, and decide to use his own credentials to quickly verify the content. In that case the attacker who controls the malicious page has now access to the analyst account.

Another possible OPSEC failure is **cross contamination**. It happens when private data of one party leaks to another. It can happen in different scenarios. For example, an analyst might be investigating a phishing webpage. When investigation is completed and they switch to handle different tasks (e.g. login to their work email), in doing so they might mistake the webpage that was still open during the investigation and provide their credentials to the phishing webpage. It might also happen that password managers autofill-in phishing webpage with credentials of the analyst.

Execution of a malware on a local machine. It might seem an unlikely situation as there are multiple controls in place by default on a SOC analyst's machine. However, it might happen that an url which links to a download page is open by the analysts on his local machine and the file is automatically downloaded by the browser.

Indirect risk for the customer

Disclose an incident to the public. When an incident happens other than damage due to the cyber attack the company might suffer reputational damage. It is therefore important that the analysts do not accidentally inform the public that a certain company has been breached. Especially before the company has started a recovery plan, which involves informing the public about the incident. Often the SOCs use codenames when refer to their customers so in the case the investigation becomes public for some reason, the identity of the customer is still protected.

Cross contamination. Is an OPSEC failure which affects also the customer, and not only the investigation itself. It might happen that an analyst is handling multiple investigations at the same time, and it so happens that an incident is escalated to the wrong contact person leaking highly confidential information to another company. This is especially dangerous because during the escalations the name of the company is usually used, as the codenames are for internal use.

Leaking of sensible/confidential information. When the analysts examine the files looking for clues of an attack, some of these files might contain confidentials information. One of the common practices among the analysts is uploading malware samples to online services such as VirusTotal to check if it is malicious. However, the file uploaded to their servers remains available to the public, and therefore if one of these files contains sensitive information for the company it would leak this content.

It is important to note that even though there might be other obvious elements that rightfully belong to this list this was the result of cognitive task analysis validated from experts in the sector. Therefore additional OPSEC failures which have not been mentioned by the interviewee are not part of this list. Additionally, other actions of the analysts which can have direct catastrophic consequences for the company, such as taking down critical pieces of infrastructure during the investigation (e.g. shutdown a server), are not included as they are indeed mistakes of the analysts but not OPSEC mistakes.

4.3.3 Causes of failures

Many reasons can be the cause of the presented OPSEC failures. We will try to describe the reasons this might happen from a cognitive perspective. First of all **lack of awareness**, this reason is the root cause of most of the OPSEC failures, it is a broad concept but can be broken down into two pieces. Not being aware about the capabilities and inner working of the tools used. Not knowing best practices to use, this can be caused by lack of experience or simply by a lack of a source to draw on for information about OPSEC best practices. This research should contribute to mitigate the latter case. The second part of the lack of awareness regards not knowing what's are the inner workings of the tools used to perform the investigation. It is important to be aware of what happens behind the scene, for instance, if a specific malware analysis tool establishes connections to IP addresses hardcoded the analyst must know so that it is possible to take countermeasures.

Second is the **underestimation of attacker capabilities**. In order to detect many of the OPSEC failures listed, an attacker would require additional work and infrastructure (for instance to monitor unwanted IP connections). However, an highly motivated attacker can find the investment in this additional work worth it. An analyst could underestimate the attacker's motivation, but also the tools that are available to the attacker. It is a safe assumption that an attacker has easy access to the same tools an analyst has (with the exception if those tools are custom made). Such an assumption will also be validated in phase II of this research where with simple open-source tools available to everyone, we will demonstrate that it is possible to set up a simple SOC.

The third element, **distraction**, can cause problems like cross contamination and indirect risks for the customer such as: accidental disclosure of an incident to the public, or the leaking of sensible information by uploading sensible material to public services. This element is highly correlated to another big issue amongst analysts. The so called burnout[4].

Lastly, **overconfidence** even though might be rare is a cause of OPSEC failures. An analyst who is overconfident of having caught the bad guy in time might not follow an assume-breach escalation path and communicate to the customer through compromised mediums.

[4]https://swimlane.com/blog/analyst-burnout-signs

4.4 Summary

This chapter described how the first phase of the research was carried out. Different Cognitive Task Analysis techniques have been used to study the behaviour of a SOC analyst in different scenarios. The use of CTA turned out to be quite effective in eliciting analyst tacit knowledge, especially when performed from the perspective of an “insider” trained as tier1 analyst. The result of the semi-structured interview revealed the typical workflow of the analysts and highlighted the important aspects that influence them during an investigation. It also provided useful insights on the training process of the analysts. One of the objectives was to discover “where” and “how” analysts learn to preserve their OPSEC. An important result was discovering that such practices are mostly learned in two ways: by shadowing more experienced analysts, or by learning from mistakes. Both the options are not ideal, and the field would benefit from a more structured way to transfer this knowledge. The combination of semi-structured interviews together with hypothetical scenarios analysis resulted in fruitful insights from the subjects. Analyse the scenarios two times, the first in a regular way and the second time by critiquing an hypothetical analyst lead a surprisingly good response from the analysts. The think aloud critique of analysts lead to identification of three major categories for OPSEC failures: “compromise investigation”, “risk for the analyst” and “indirect risk for the customer”. Finally, the cause of such failures have been analyzed and discussed.

Chapter 5

Wargame

The previous chapter presented the findings of Phase I of the researcher, showing how the analysts can compromise their investigation and hence the security of their customer by committing OPSEC mistakes. The following chapter will instead show how the impact of analyst's actions can influence the Red Team. In doing so first, the planning and the organization of the wargame experiment will be discussed. The general requirement and design choices will be introduced. We will continue then with the details of the technology used and elements included in the design of the scenario. The challenges encountered will also be presented. Finally, it will be described how the findings of the previous chapter will be validated by modelling a Blue Team variable used in the experiment.

5.1 Cyber range

The previous chapter presented the findings of Phase I of the researcher, One of the goals of this research was to create an environment which will enable future research in the field of Red and Blue Teams interplay. As discussed in the previous chapters, one of the main obstacles for researches that target Red and Blue Teams is to find a way to observe them in a ”natural” environment, i.e. during a Red team assessment. The solution to this problem was to organize a cyber event where it would be possible to observe the natural behaviour of the two teams. One of the benefits of observing the teams interacting in a virtual environment is the possibility to recreate specific scenarios, and for instance, being able to study how the same actors behave with different stimuli.

The use of cyber events to study cyber behaviour is not a new practice. However, such events are usually big events which require the participation of many people and the use of many resources. One of many examples is Locked Shields. It is organised every year by CCDEOE Some examples. Every year CCDCOE [1] organise **Locked Shields**, an exercise where cybersecurity experts can defend national IT systems and critical infrastructures under real-time attacks. It consists of Red team vs. Blue Team exercise, where the latter are formed by member nations of CCDCOE. It is a rather big event with about 4000 virtualised systems that the participants have to attack and defend. It includes every aspect of a cyber crisis, from the attack, incident reporting, executing strategic decisions, solving forensics up to handling legal and media challenges [81].

One of the goals of this Phase Is to create a reusable environment to host wargame events, which can overcome the current limitation of studying cybersecurity dynamics during big events. There are existing researches which tried to achieve similar results [82] [83] [57]. However, as it was observed by Yamin et al [84], the researcher is always a passive observer. Based on the examined literature, no attempts at involving the researcher proactively during the experiment has been attempted before. This project will differentiate from the existing research by introducing a state of the art built-in monitoring system which the researcher can use to emulate Blue Team actions and interact with the subjects of the experiment directly. Thanks to the previous phases of this research, the researcher has also acquired the necessary skills to perform this step.

5.1.1 Requirements

The first step is to identify the requirements for such an environment. MITRE identified five factors that determine the success of cyber wargame [70]. The first is the cost, due to the significant investment of time, effort and technologies which are required to set it up. *Realism and fidelity*, the planning of a wargame should take into consideration has an adequate level of details. *Scenario preparation*, it is important that the additional information and context provided to the participant is plausible, otherwise, they might take their mind out of the game. *Attacker preparation*, a real attack would last weeks or months, it is essential to furnish the attacker ways to exploit the system within the constraints of the wargame. *Knowledgeable players*, the

[1] NATO CCD COE, officially the NATO Cooperative Cyber Defence Centre of Excellence is one of NATO Centres of Excellence (COEs), for training on technically sophisticated aspects of NATO operations

quality and validity of the insights obtained from the wargame depends on the capabilities of the players. All the previous factors have been taken into consideration and have influenced the preparation of the cyber wargame.

Additionally, there are other requirements. It has *built-in monitoring capabilities* in order to allow the defensive team to monitor and respond to cyber threats appropriately. It should be *flexible* enough to support multiple network topologies and systems and therefore allow it to adapt to a range of different scenarios and attack paths. *Open-source* is an obvious requirement as it will also support the scientific community in fostering this research. Finally, it should be fully *automatable*.

Rice and Edgar justify the need for such requirements in their research "Experiment as a service". They identified that the main problem in the cybersecurity research field is the lack of rigorous experimentation which is mainly influenced by two factors: the lack of repeatable and reproducible experimental environment, and the lack of realistic models to recreate enterprise systems [82]. A Realistic and reproducible experimental environment can be redeployed and verified by other academics; it was also pointed out that realism should encompass business processes as well as simulated users.

5.1.2 Design Choices

The previous section described which are the requirements for the wargame environment and why they are necessary. Following will discuss the specific design choices and how they will help meet the requirements.

Active Directory Environment Based on a preliminary interview with some members of the Red Team, it was discovered that in many of the assessments they perform the customers have a windows domain network supported by Active Directory. Further research revealed that Active Directory is used by most of the big organization worldwide. For these reasons, it has been decided to design an active directory environment. This choice will help to meet the realism requirement.

Public Cloud Provider In order to satisfy the flexibility requirement, It has been decided to take advantage of cloud providers to host the organization's network entirely and to use infrastructure as code tools to create an easily deployable and

reusable environment to satisfy the automatization requirement. This choice also satisfies the cost requirement since cloud providers nowadays are providing cheaper and cheaper hosting capabilities. There are multiple reasons for this decision. The first reason is that the creation and configuration of a virtual environment which also includes planned vulnerable paths is time-consuming, and in case one of the elements of the environment stops working it should be possible to bring it back to the original state fast reducing downtime during the event. The second reason is that in order to verify the data and the conclusions of the experiment by future researchers, it is necessary to recreate an environment with precisely the same characteristics. The final reason is that this work is one of the few that explore the interplay of Red and Blue Teams, and for this reason, one of the goals is to create a flexible tool that can be used and expanded by future researchers who want to research this topic further.

It is worth mentioning that other researchers addressed the issue of deploying a virtual environment which can be used to validate and test experimental results in the cybersecurity field [82] [83]. However, the choice for technology used differs substantially from the choice made in this research. The main difference is in the use of OpenStack [83] to manage the infrastructure.

Infrastructure as code IaC is a new paradigm which refers to the process of managing and provisioning computer data centers through machine-readable definition files, rather than physical hardware configuration or interactive configuration tools. [85]

Monitoring capabilities In order to be able to capture the interplay between Blue and Red Team fully the technologies available to the Blue Team must meet the standard of the industry. For this reason, part of the interviews performed in Phase I of the research has been used to identify possible technologies which can be used to create a realistic Security Operation Center (always keeping in mind the cost requirement, as some of these technologies can be rather expensive).

Vulnerabilities Another critical design choice is in the vulnerabilities which were introduced to support the attack paths. The choice of the vulnerabilities was lead by two baselines. First, it should be present in the scenarios which have been used to identify OPSEC failures and to model the behaviour of the Blue Team in the

previous chapter. Second, in order to enable the Red Team's action to be as realistic as possible, the vulnerabilities should cover not only technical components but also the process, and the people involved (as Rice and Edgar [82] highlighted it).

5.1.3 Technology Used

Cloud provider It was a common choice amongst researchers to use OpenStack as the underlying technology to manage the infrastructure. Openstack is OpenStack is a free open standard cloud computing platform that supports both public and private cloud. Being an open-source project, this seems an ideal choice in many situations, especially in the academic field. However, other public cloud providers come with builtin solutions which are the de facto standard in many cybersecurity situations. An example is Microsoft Azure, among the other services Azure provides state of the art SIEM (Sentinel), support for Windows machine and Active Directory environment. It is also allowed to easily integrate other services such as Office365, which is largely adopted by most of the big organizations. Other public cloud providers have been taken into consideration to host the virtual infrastructure: Azure, AWS, IBM, Google. However, when it comes to creating a Windows environment, Microsoft Azure offers more support. Moreover, the native SIEM solution (Azure Sentinel) that it offers has been decisive in the choice of using Azure.

In summary, the transparency of an Open source project (such as OpenStack) already widely adopted by the community has been sacrificed in favour of a proprietary solution (Azure) in order to be able to create a realistic environment more easily.

Terrafrom The whole environment has been created following the "Infrastructure as Code" paradigm. Terraform is an open-source tool that uses a declarative language to define the elements of a virtual infrastructure as well as how they are connected together. It abstracts all the logic needed to deal with cloud providers and allows the end-users to reuse the same code to deploy their infrastructure on different providers. Once the provisioning of the elements is complete, it is necessary to configure them. However, it offers little control over the configuration of the system and the software installed. For this reason, it has been decided to couple Terraform with Ansible.

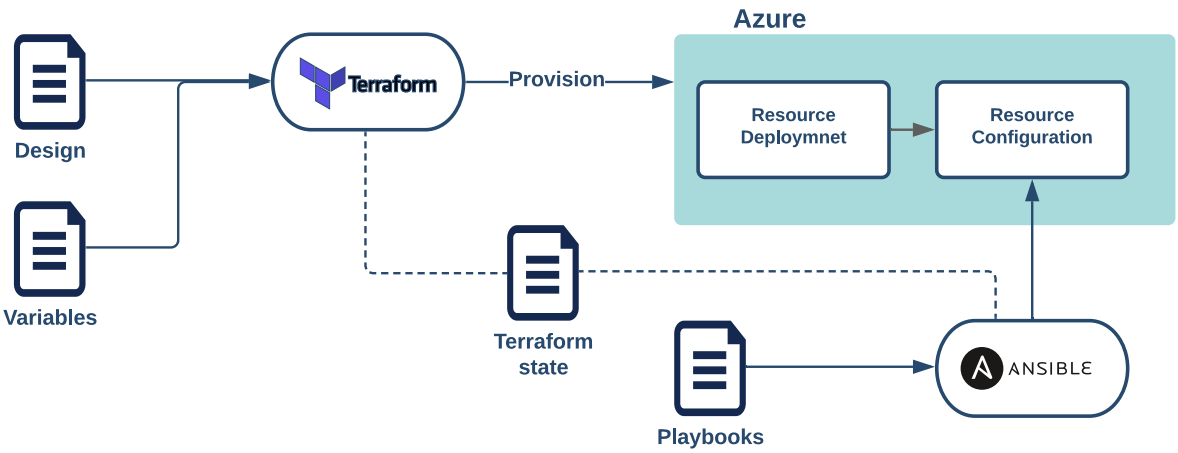


Figure 5.1: Infrastructure Deployment

Ansible is a configuration management tool, which is used to manage and configure servers. Ansible works by defining a certain desirable state for the servers, and then it enforces such a state (install packages, running services, configure software, etc.). It uses a declarative language to create configuration files called playbooks. Note that Ansible does not describe ”how” to achieve a specific state but ”what” state to achieve. In this way, it abstracts the details and the steps required to install and configure a particular computer, and is; therefore, the developer can focus on the design of the infrastructure itself.

5.1.4 Deployment

The following section describes how the described technologies come together to achieve the requirements. The scenarios are composed of a series of files that together will create the environment. Terraform files are used to define which systems are present, their high-level roles (e.g. Workstation, Domain controller, Web Server), and the network topology. Each Terraform scenario can then be customized even more by changing variables (e.g. whitelisted IP addresses, number of workstations). Terraform then proceeds with the provisioning phase, where based on the cloud provider which has been chosen APIs will be invoked to deploy the actual components. In the process, Terraform will write the changes and the state of the resources in a Terraform state file on the local machine. Subsequently, the configuration phase begins, and Ansible uses playbooks to change the configurations of the machines deployed in the cloud (see figure 5.1).

Modules A series of basic modules have been created, which can be combined to form the final scenarios. For example, the Web server module is configured to deploy a Windows machine, update it, install the IIS web server, load a premade website, and join the domain. Another example is the Domain Controller module first deploy the windows machine and install the necessary packages, and then the domain is configured, the active directory structure is created with organizational units and populated with users according to what has been defined in the scenario. Each module accepts an Ansible playbook as a parameter. Therefore it possible create a different set of vulnerabilities based on the requirements of the scenarios just by creating different playbooks. Terraform and Ansible are easily parallelizable; consequently, the whole operation takes a matter of minutes[2].

Following are listed the modules which have been developed, and that can be combined together to build more complicated scenarios: domain-controller, database, exchange-server, workstation, server, network. There are also a number of playbooks which can satisfy different needs. At the moment of the experiment, the task needed to complete the development of the scenarios, however, still required manual some manual configuration (e.g. Exchange server). The reason is that the amount of work required to automate some of these steps exceed the time available for this research, however, this does not imply that the task is not achievable, the issue would be easily solved with the aid of more researchers.

5.1.5 Design of the Scenario

The modules described above have been used to create the environment of the scenario for the wargame. The following section will present the main elements included in the virtual environment. As described in the methodology chapter 3, the scenario for the wargame is designed starting from the hypothetical scenarios analyzed in Phase I. This choice also reflects on the choice of infrastructure elements that will be included in the final environment. Figure 6.1 gives a high-level overview of how the elements of the infrastructure fit together. A detailed view of every element can be found in appendix B.

[2]The process of deployment can take longer in the case of the configuration of complex services such as an Exchange Server

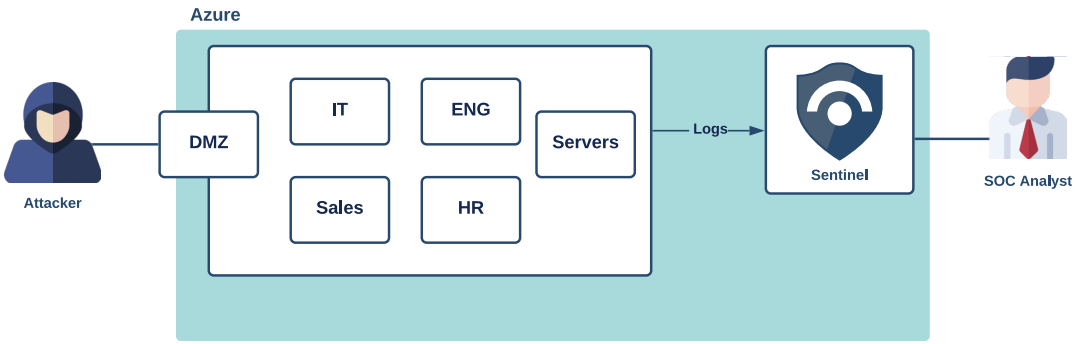


Figure 5.2: High level overview of the infrastructure

Infrastructure

Subnets The virtual environment has five sub-networks. The first network is the DMZ, where the webserver and the RDS gateway are located. The second network is the Internal servers LAN, in this subnets resides the file server and the internal server with the company database. The other four subnets host the workstations of each department(IT, Sales, Engineering, and HR).

Firewalls Basic network segmentation has been included. The DMZ allows connections to the internal network only from the RDS gateway to the fileserver. The internal server hosts a web app used by the sales department and is therefore reachable only by the sales subnet. The workstations from the IT department can access every other subnet.

Servers A web server hosts the company website. An RDS gateway located in the DMZ provides remote desktop capabilities. A file server is used to host files and share them in the network. An internal server hosts web applications for internal use and is connected to a database which contains company critical information. A domain controller is used to manage the whole active directory infrastructure.

Mailboxes The environment includes a Microsoft Exchange server. Such a server is used to provide mailbox capabilities to the users of the domain. Every user has been assigned a mailbox which can be used by the Red Team to contact the user a build phishing campaigns.

SOC

One of the requirements is the presence of builtin-monitoring capabilities to create a SOC integrated into the virtual environment. The starting point for this feature was an existing open-source project called SentinelAttack [86], which includes configuration files to set up a SIEM solution (Azure Sentinel) and a workbook to visualize data collected from the data sources.

Log Collection Each virtual machine has been configured with Sysmon. System Monitor (Sysmon) is a Windows system service and device driver that, once installed on a system, remains resident across system reboots to monitor and log system activity to the Windows event log. This tool provides great visibility on the actions the Red Team performs as it provides detailed information about process creations, network connections, and changes to file creation time [87]. Being Sysmon such a powerful tool it can collect every kind of event, but the downside of it is that the analysis of such a large amount of logs can become problematic. For this reason, Symon has been configured to map each event to an ATT&CK framework technique. It then sents only the events that are likely related to an attacker's action. The configuration used maps the events to 159 techniques of the ATT&CK framework, allowing to attribute actions to Red Team easily.

SIEM The SIEM solution used to support Blue Team capabilities during the wargame is Azure Sentinel. Microsoft Azure Sentinel is a scalable, cloud-native, security information event management (SIEM) and security orchestration automated response (SOAR) solution which includes alert detection and threat response functionalities. Sentinel offers many functionalities; however, not all of them have been used during the experiment. Some of this functionalities include: connection of multiple data sources (Sysmon was the only used), analytics correlates alerts into incident-based, machine learning rules to identify anomalous behaviour in the network, deep investigation tools and threat hunting tools. Since all these functionalities are part of Sentinel, those are easily accessible to the analysts operating in the wargame environment if they need it.

SOC Console Sentinel uses the workbooks to take the logs collected from Sysmon, enriched and mapped to the corresponding ATT&CK technique and displayed on the SOC console. It is then possible for the analyst to dig through the logs to

identify suspicious actions. This first version of the wargame environment still does not include a function to generate alarms based on some rules. Nevertheless, due to the lack of white noise, it is not considered a big limitation, as most of the events detected should be related (to some degree) to the actions of the Red Team. More details on the views from the SOC console and the capabilities that are offered can be found in appendix B

Attack paths

In the previous chapter, the interviews with the SOC have been used for two different purposes. The first purpose was to create a list of OPSEC failures that can reveal the status of the SOC analyst investigations. The second purpose was to support the decision of the attack paths to be used. In order to ensure that the attacker can eventually reach the crown jewels, a series of attack paths have been planned and implemented. The last step in the design of the scenario is the creation of these attack paths. The following section will give a high-level overview of the attack paths, a more detailed explanation can be found in appendix B

The attack path, just like the hypothetical scenarios, used in Phase I are designed accordingly to the steps in the Unified Cyber Kill Chain. Feedback from the interviewed analysts regarding the realism of the scenario presented was also taken into consideration when planning the attack paths. Each of the scenarios has been included in the attack paths, this way the realism is ensured, besides it also guarantees that the researcher is able to approximate the actions of a real analyst very closely thanks to the insights of Phase I.

The goal of the Red Team was to reach two crown jewels: the first one was "*gain access to the customer database situated in an isolated network*". The second one was "*became domain admin*". The initial step to obtaining both the crown jewels required the subject to perform a phishing campaign to obtain the initial foothold and the escalate privileges in order to be able to perform lateral movement. After that to obtain the first crown jewel, it was required to go through the Sales department. The path to reach the second one required to obtain access to the HR department first. A series of sub attack paths that could be linked together to reach the crown jewels are listed below:

1. Initial foothold path. The participant uses the information on the website to mount a phishing attack, obtaining credentials or a reverse shell on the remote

desktop session of a member of the engineer department

2. Lateral Movement 1 from Engineering. The participant can perform a password spray attack against the members of the IT department. Some of the employees have weak passwords.

3. Lateral Movement 2 from Engineering. The participant can use user hunting to perform a derivative local admin attack gaining access to the Sales department

4. From sales to customer DB. The internal web server hosts a web app which is accessible only from the Sales or IT subnet. The web app is vulnerable to SQL injection attacks. The participant can

5. From department admin to HR. Each department had an admin who operated through an admin workstation. Such workstations were configured with unconstrained delegation flags set to true(explain unconstrained delegation). The participant could use internal phishing to

6. Kerberoasting SQL service. The participant could perform a kerberoast attack obtaining the credentials

All the described attack paths could be revealed to the participant by running a Bloodhound ingestor and subsequently by analysing the graph generated using Bloodhound[3].

5.1.6 Independent variable

One of the main challenges in the cybersecurity research field is the difficulty to establish scientific validity from other researchers' results. The reason is that reproducing the same results is not always possible due to the high number of uncontrolled variables during cyber events. For instance, the network configuration or the specific vulnerabilities will change in different iterations of the events. The reason is that these events are not organized with the primary objective to support specific research experiments, whereas they are organized to engage and challenge the participant. Such events are a great source of data and insights, but the downside is that they do not provide a reliable testing environment. Rice and Edgar discussed

[3]BloodHound is a tools which uses graph theory to reveal the hidden relationships within an Active Directory environment. Attackers can use BloodHound to easily identify highly complex attack paths that would otherwise be impossible to quickly identify.

this issue in their work ”Experiment as a Service” [82], they highlighted the fact that ”the number of uncontrolled variables is often too great to draw conclusions about repeated or comparative experiments”. For this reason, in an attempt to reduce the number of uncontrolled variables, the Blue Team's actions have defined as a controlled variable, and the regular users have been set as a constant. The reason for considering regular users a constant is that their actions are scripted and repeated independently from the Red Team actions. The following section will describe how the Blue Team variable has been defined based on the finding of Phase I.

Blue Team actions

The independent variable used during the experiment consists of a set of actions the analyst can do during the investigation. The actions have been divided into two subsets. The first subset consists of the basic activities performed during a ”regular” investigation. Such activities have been identified first of all by analysing the existing literature, and then by observing the analysts work identifying a reproducible pattern. This subset of activities can be considered as the base reference; variations on activities performed by the Blue Team can be used to study different reactions of the attacker. The second subset can be seen as some variation on the base set. It consists of the OPSEC failures identified in Phase I. The actions stray from the typical behaviour of a Blue Teamer, and may or may not influence the Red Team, which is the purpose of this research.

For ease of use and reference during the experiment, the first set of actions have been represented as a scheme. The resulting scheme is a modified version of the general analyst's workflow presented by d'Amico et al [71]. On top of that for each scenario, the specific actions described by the analysts in Phase I will be reproduced. During the experiment the steps described in the schema (figure 5.4) are followed, however the decisions are based on the experience of the analyst and vary from case to case. Three steps have been highlighted were modification to the actions of the analysts can be applied to observe the response of the Red Team, i.e. “data inspection”, “integration with external sources” and “responsive action”. The OPSEC failures as in Phase I are applied there.

Figure 6.1 shows some of the elements that determine the state of the wargame. As will be discussed in the next section the behaviour of the regular users is imposed by a series of “routines”. The cyber attacker behave freely performing cyber attacks. Whereas, the Blue Team's actions are determined by the workflow (see fig-

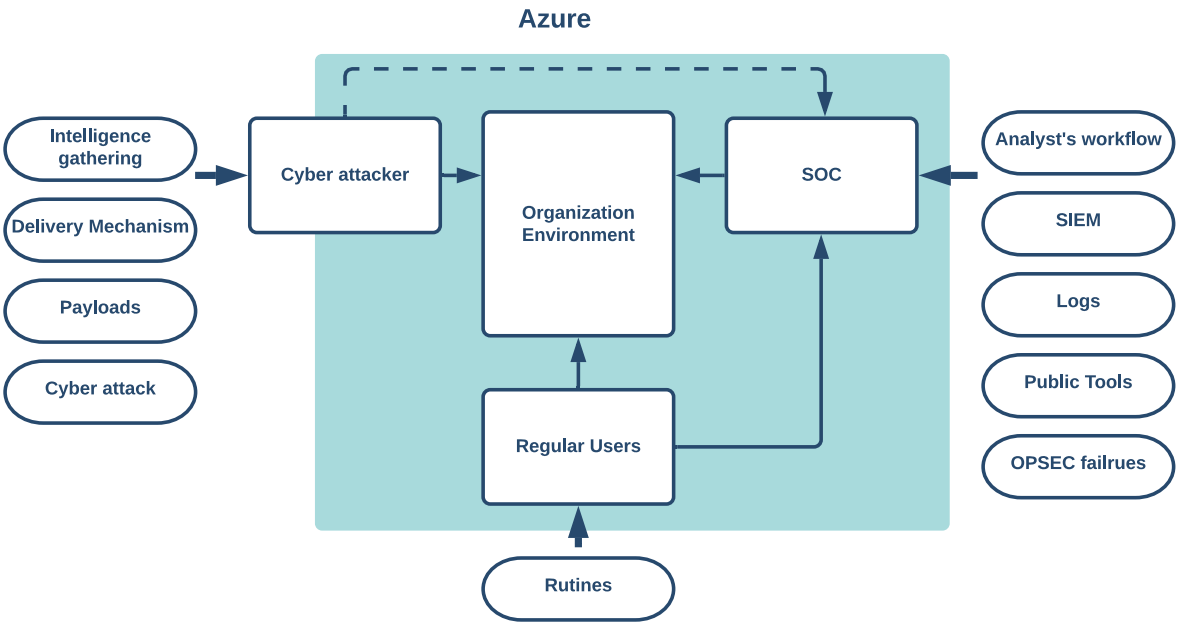


Figure 5.3: Interaction between the BT controlled variable and RT uncontrolled variable in the virtual environment

ure 5.4) but also influenced by what logs they can see through the SIEM, and from the result of the scans performed with public tools. Moreover, Blue Team actions are determined by the OPSEC failures that are being evaluated.

An important decision that characterizes this research was not to include Blue Team members during the envent. The simple reason for this choice is that other participants would have required to observe and record their behaviour as well, and would have reduced the control over the Blue Team's actions. Besides, according to a previous researcher that studied Red and Blue Teams, correlate action and response of two teams simultaneously is a remarkable challenge [55]. We choose to simplify the research to obtain more clear-cut results.

Regular users

As discussed in the introduction chapter, a Red Team does not exploit only technical vulnerabilities, but also people and processes. In order to make a complete and truthful observation of the Red Team assessment, it would be necessary to include regular users as participants in the wargame, or at least simulate the actions of such users. During the wargame, a set of ”routines” have been defined and periodically executed by the researcher. Such ”routines” represent regular users. For example,

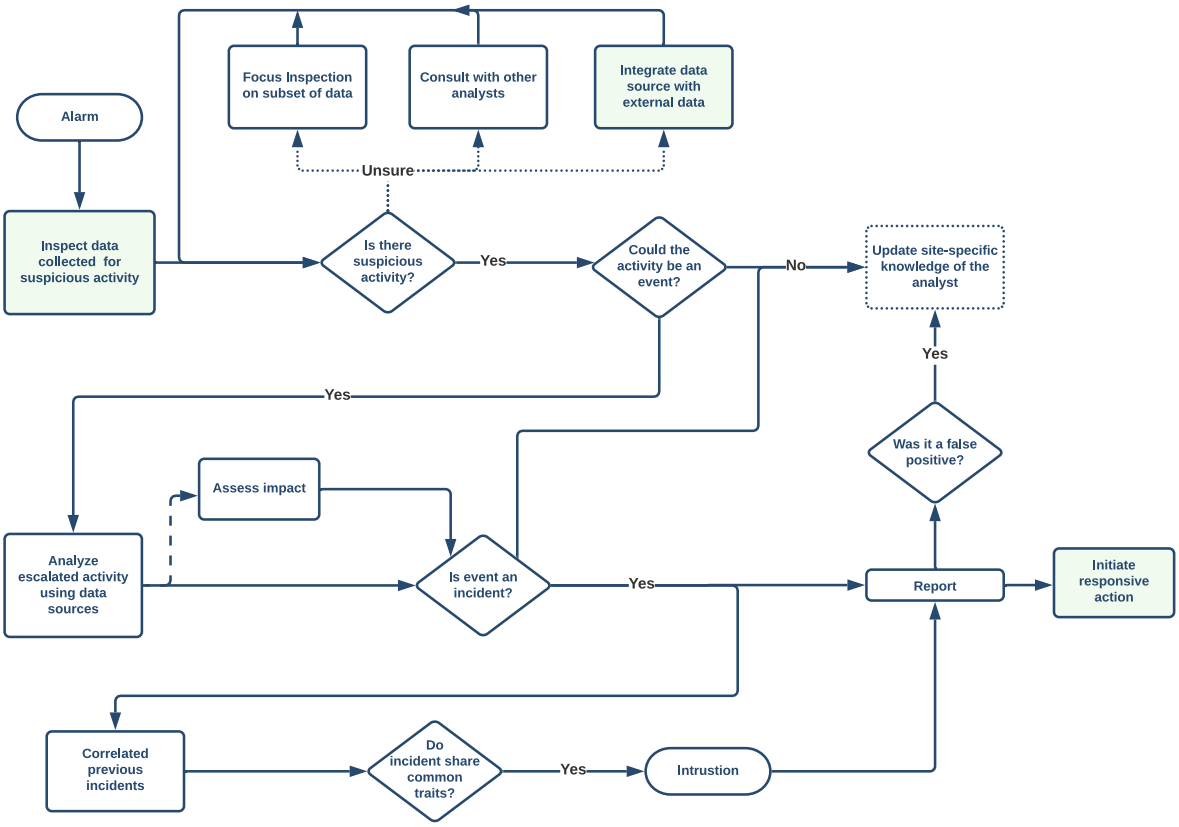


Figure 5.4: Analyst workflow

opening a remote desktop session into each workstation and interacting with the Red Team. The main actions that have been simulated were two. The first was opening of emails and eventually falling for phishing attempts by entering credentials or executing malware that was hidden in the email. The second action was the periodical opening of shared folders on other users computers. The first action simulates the interaction user-outside word, and the second action simulates the interaction user-user inside the network. This set of regular users action can be used to either obtain the initial foothold in the system or to perform lateral movement (as in the case of attack path 3). All these actions can be changed if needed in future research. However, it is advisable to reduce the variations to the minimum in order to be able to obtain more scientifically verifiable results.

To conclude some considerations regarding how the Blue Team actions can be contextualized. The main challenge faced by analysts is the process of sensemaking [65], and there is much uncertainty associated with this process. An analyst can not observe the actions of the Red Team directly, but only through the traces they leave on the system, such traces are often ambiguous. Moreover, analysts can only observe the tip of the iceberg, the traces that they can observe accounts only for a small part of the Red team's behaviour. Besides, meaningful indicators of attack

activities are often very scattered [55]. Such factors and a high level of uncertainty makes it so that it is not possible to have a static list of actions which defines the Blue Team variables, whereas a flowchart that can be used to determine the next step in the investigation is more suitable in such a dynamic environment.

5.1.7 Limitations

The following section will discuss the limitation of the environment and of the scenario created for the experiment. The first and main limitation is the **lack of "white noise"** in the network. In a real environment every system would be used by real users, and this in turn would hide the actions of Red Team to the SOC analyst. Additionally the presence of real users can open new attack paths for the Red Team and improve the realism. This problem was partially overcome by defining a set of routines for the regular users in the experiment, however the amount of noise generated is not sufficient to achieve the desired level of realism.

Another limitation was the **limited time frame** of the event. A regular Red Team assessment would last for days or weeks, whereas the wargame lasted only one day. For this reason some of the most time consuming parts of a Red Teaming have been facilitated. Reconnaissance and obtaining the initial foothold is the part that requires the Red Team to invest a lot of time in figuring out scenarios the user would fall for, sometimes even by creating trust by exchanging multiple emails over time before trying to actually get access to the system. For this reason the amount of information available on the website was limited in order to make it immediately clear what the user would fall for. The second time consuming task is obfuscation of the malware. To avoid the team to spend too much time in trying different techniques to obfuscate the malware the antivirus was disabled on some of the workstations, moreover the malware scanning function of the mailboxes on the Exchange server were disabled. Additionally a series of milestones were defined that the team should have had reach before a certain time. If the milestone was not reached an hint would have been given to the team to take them back on track. However, this scenario did not happen and the team did not need additional hints.

A final and important limitation which heavily influenced the outcome of the experiment was the presence of **unintended attack paths**. All the attack paths have been tested by the researcher in order to verify that the system was actually exploitable. However, a mistake that has been made was the one of not performing an extensive test on the rest of the environment to identify alternative ways to exploit

the system. Iterating the experiment multiple times in the future and patching united vulnerabilities should help to mitigate the problem overtime.

5.2 Experiment

The previous section described how the environment to support the experiment was designed and implemented. As it was described in chapter 3 the last step of Phase II is to use the infrastructure and the scenarios developed to observe the influence that BT OPSEC failures have on the actions of the Red Team. The following section will describe in detailed the purpose of the experiment and how it was carried out.

Goal The general objective of the experiment was to observe how the Red Team behaves in a realistic environment when a Blue Team is actually monitoring the system and hence their actions. Furthermore, the experiment should show to what degree the Red Team is able to detect the actions of the Blue Team.

Data There are two kinds of data generated by the experiment which will be further analysed. The first are raw logs that are collected through the Sysmon system which reveals details on the actions performed by the Red Team . The second is a qualitative kind of data that comes from a series of short interviews with the Red Team after the event.

5.2.1 Wargame Day

Description of the Participants The participants in the experiment were seven experienced Red Teamers each with at least 2+ years of working experience. All the participants were part of the Northwave Red Team. They hold multiple certifications such as OSCP[4] among others. The participant were divided in teams and each team was formed by at least a senior Red Teamer.

[4]Offensive Security Certified Professional (OSCP) is an hands-on ethical hacking certification offered by Offensive Security which certified that the ethical hacker has penetration testing skills

Blue Team Actions

The following section is a log of the actions taken by the Blue Team. All the actions where informed by the first part of the research. As was discussed in the previous sections the first part of this research served two purposes. The first purpose was to identify how the Blue Team could compromise its investigation and reveal the status of their analysis to the Red Team. The second purpose was to define an independent variable for the wargame experiment. One of the many limitations identified by previous researches on Red and Blue Team interplay was the difficulty to control and analyze an environment with so many different actors [55].

Mainly because of the time constraints of the event, the Blue Team did not take any active actions to stop the progress of the Red Team. However, this aspect of the Blue Team investigation should receive more attention in future research because it is the one which has the biggest impact on Red Team operations and is the actual reason that motivates the adversary in taking countermeasures to try to detect the status of the Blue Team's investigations.

The independent variable used during the experiment can therefore be summarize as follow: *Actions of SOC analysts*. As was discovered in the preliminary research SOC analysts perform mostly semi-active investigations. They analyse logs and data "offline" as much as possible, and interact with the system only in few cases. Most importantly, SOC analysts do not take direct actions to stop the intruder, whereas they escalate the incident to the customer letting them to decide which action to take.

The variable observed during the experiment measure the actions taken by the Red Team. Therefore the function analyzed by the experiment can be describe as: *Response of the Red Team to the actions of a SOC analyst*

5.2.2 Data Collection

The data collection process is simple. Each participant was required to write a short report of their actions right after the event, describing the main steps they took to reach the crown jewels. Each report has then be compared with the logs collected by the SIEM and cross referenced with the actions performed by the researcher. Following the data collected for each team will be presented. Even though the teams

have tested different techniques at each stage, not all of them were successful to breach into the system further. In order to keep the analysis simple only the successful attack paths and the investigation activities of the SOC analysis to pursue such actions will be analysed further. Additionally, the actions performed by the analyst in response to each step of the attacker have been collected for further analysis.

Team A

Team A was the first to discover the unintended path and to reach the crown jewels. They also used interesting ”non technical” ways to obtain the initial foothold (phone call). They breach the system this way: Performed a social engineering attack through the phone. The started by performing Reconnaissance of the domain. Then they initiated a phishing campaign to obtain initial foothold. At this stage, the team found the unintented attack path. RDP to the exchange server. They performed a Lssdump of the nwsoc user credentials. Then the continue by running PsExec to the Domain controller. Finally, they established a connection to DB and full database dump. An those were the actions performed by the analyst: Finding traces of Team A in the SIEM was harder than for the other teams. The phishing email was analyzed, and the phishing domain scanned with public scan.

Team B

Team B applied a large variety of different techniques and explored many of the intended attack paths, however, they finally reached crown jewels via the unintended attack path. The chain of actions that lead to obtaining the crown jewels was the following: Reconnaissance on the company domain. They started by harvesting email addresses and began a phishing campaign. For the campaign, they used both credentials phishing and fileless attack. After obtaining the initial foothold, they run bloodhound for situational awareness. They concluded their attack by performing a Lssas dump and obtained domain admin credentials (via unintended attack path). In response to the actions of the Red Team the Blue Team did the following: performed analysis of phishing email, performed a public scan of the phishing domain, run malware in a sandboxed environment without an internet connection, perform connection and Nmap scan of the Team’s infrastructure.

Team C

Team C was the one that followed the intended path for longer before deviating to the unintended path only on the last steps. They also used an interesting technique to better exploit the human factor. They first sent an email to prepare the ground, and subsequently, they proceeded with the actual phishing email. Therefore the victim was more likely to fall for the attack. The team first begin with a Reconnaissance phase and a phishing campaign, they chose the fileless attack vector, run bloodhound after getting an initial foothold, perform password spray and performed lateral movement to the IT department, at this point the team dumped the domain admin credentials and followed the unintended path to obtain the crown jewels. In response to their actions the Blue Team: perform an analysis of the phishing email, a public scan of the phishing domain, once the malware agent was detected it was analysed in sandboxed environment without internet access, and scanned the domain with Nmap.

Comment on the wargame day

Unfortunately, the Red Team used their custom malware during the attack. For this reason, it has been decided to not perform all the OPSEC mistakes that were planned. In this way, we avoided compromising the future operations of the team. Suppose the malware was uploaded to public service such as VirusTotal or any.run[5] as it was planned it would have resulted in increasing the detection rate of the Red Team in future engagements. It was instead decided to ask during the follow-up interviews a "what-if" question, to evaluate if the team would have been able to catch such activity.

Follow-up Interview

Based on the actions taken by the Red Team during the wargame there were no major indicators that the actions of the analyst triggered any anomaly in the behaviour of the attackers.

Due to the unforeseen strain of the teams, the researcher did not investigate the

[5]Cloud-based malware analysis service which is used to run suspicious executables or visit websites, and records system and network-level activity

actions performed by the teams after the deviation. The researcher was actively hunting for possible events to investigate focusing the attention on elements of the infrastructure that were actually meant to be exploited. For this reason, the analysis will be framed solely to the actions the Red Team performed while they were on the "right" paths and that were actually investigated by the Blue Team. Eventually, all the teams deviated from the intended path, however they all went through the same initial steps to get the foothold on the system.

In order to answer the third research question, the teams were briefly interviewed after the event to have a feedback on the actions of the Blue Team that they might have observed.

Did you monitor your infrastructure for suspicious connections?

Would you have been able to detect if malware was run in a sandbox environment?

Did you monitor public websites for phishing email?

Did you monitor public services to verify if the malware was being uploaded?

To what extent the presence of SOC analysts influenced your actions?

Did you observe any suspicious event indicating activities of the Blue Team?

The members of the Red Team interviewed commented that the major reason why they did not change their strategies was that the Blue Team was not perceived as threat. In their viewed the experiment was perceived as a game, hence they did not adopted many of the techniques they would have otherwise used during a real assessment with an opposing Blue Team. For future iterations of the wargame if the Rules of Engagements clarify the position of the Blue Team on this point they will act differently. For instance by obfuscating malware, planning spear phishing[6] campaign opposed to regular phishing campaigns which are harder to detect.

We found that the team was not monitoring public services for traces of malware being uploaded or public scan being used to scan their phishing website and therefore they would not have been able to promptly react to the event. However, in case the phishing website was flagged as malicious or if the malware was detected by only AV services they would aspect to start seeing their phishing emails or malware

[6]Spear phishing is type of phishing were the email specifically crafted for one individual

being blocked. Therefore in an indirect way, they would have been able to detect such an event.

Another factor that should be considered is the fact that the team expect that their phishing campaign will be detected and malware eventually flagged, however usually the time frame between first detection and countermeasures put in place is large enough that they can obtain a foothold in the system and cover their tracks. Additionally, often the countermeasures are sometimes ignored by the user, for instance in case the phishing campaign is detected there are users that keep fall for the phish because they disregard countermeasures that the company could have taken (such as email sent to alert the employee of the ongoing campaign).

Finally, the interviewee pointed out that if their malware ran in a sandboxed environment they would have likely be able to detect it very quickly due to often unrealistic sandboxes that are used by Blue Teams. Common indicators of unrealistic sandboxes is the amount of time the malware runs, an investigation lasts a few minutes, if their reverse shell is closed soon after being executed it indicates someone might have tested the sample in a sandbox. Another indicator is if the host do not match their expectation, for instance they usually expect they malware to be run in domain joined workstation.

A fact worth mentioning is that all the teams independently decided to exploit the unintended attack path. This means that if the experiment is reproduced in the future likely other teams of attackers will follow the same path. If this is true this offers insights on how the Red Team choose which attack path to follow. Comparing the unintended attack path to the one that was planned it is clear that the former requires fewer steps to obtain the crown jewels.

Conclusion

Concluding this part of the research, there were no clear indicators that the actions of the SOC analyst had an impact on the strategy of the Red Team. However, it was confirmed during the interviews that mere presence of a SOC would have made the attacker adopt a more “safe” behaviour trying to remain undetected for longer and using more sophisticated techniques fly under the radar. However, they perceived our experiment as ”game” or a competition and not much as real Red Team assessment for this reason they did not many of the techniques they use in a regular engagement. Based on the reports, it was not possible to evince any

major change in the strategy as well. On the positive side, the experiment confirmed that an investigation carried out properly is virtually undetectable by a malicious actor. Nevertheless, OPSEC failures are still a possible risk source for Blue Team operations, and if in the future a practical way to detect such failures will exist it might highly impact the balance between attacker and defenders. It is important to get ahead of the problem and include more structured OPSEC coaching in the training of the Blue Team.

5.2.3 Lesson Learned

As was discussed above not all the expectations set for the experiment were met. Two in particular lead to sub-optimal results. The first one is the unintended attack paths that was exploited by the Red Team, and the second one is that not all of the planned OPSEC mistakes were made. This two elements will now be discussed.

As it was introduced in section 2.1.1 cyber attackers are naturally out of the box thinker. Leaving them loose to operate without many restrictions in a realistic environment resulted in the identification of attack paths that were not planned and therefore not monitored by the SOC. The first issue was the lack of **extensive tests** on the environment, however even if the environment had been extensively tested it would not have been sufficient to ensure that the attacker followed only the intended path. If building a completely secure system is a challenging task, building a system purposely vulnerable is even harder because every vulnerability introduced can bring additional unforeseen vulnerabilities. Another way to restrain the creativity of the Red Team is to define **stricter rules of engagement**, however, in doing so realism and the likelihood of successfully exploit the system is reduced.

Changing the paradigm used to deploy the infrastructure is definitely a way to reduce unintended attack paths and to improve realism at the same time. So far the infrastructure was deployed from an external perspective. **Deploying the infrastructure from inside out** means that first the IT department is created, and then the terraform/ansible script are run from that machines creating the environment similarly as in a real situation. The unintended attack path was present because the environment was build from the outside, and remenant of this remained on each machine which was then exploited by the attackers.

Another issue is that simultaneous attacks have a negative impact on both the realism and on the observations. Some of the team took advantage of artefact left

behind by other teams to escalate privileges (e.g. a fake account added to the AD environment) without knowing it not being part of the scenario. From the analyst perspective tracing action of different teams and correlating the right events was an additional challenge. Moreover, a thorough analysis requires some time and with multiple attacks going on it is easy to fall behind with the queue of events to investigate. This problem could be easily solved by performing the experiment with a single team at the time, or by including **more workforce** to the Blue Team.

5.3 Summary

The Blue Team ability to perceive the actions of the Red Team are limited to what the system is configured to monitor. In large systems, automated rules that generate alerts when suspicious activities are detected help the analyst. The experiment did not include any of these automated rules and therefore, the activity of the Red Team outside the intended path remained unobserved. The efficiency of a SOC is directly dependent on its monitoring system cover. The simple SOC implemented in this experiment was not sufficient to effectively detect a simultaneous attack from multiple actors, however, it should have demonstrated how the SOC infrastructure developed for the experiment can be easily expanded in future iterations.

The Red Team, on the other hand, hardly has the means to detect a proper SOC investigation. As it was demonstrated, all “regular” actions of the Blue Team were almost impossible to observe by the Red Team because there was no exchange of information with the outside world (e.g. inspecting logs). On the other hand, in case a mistake was made during the investigation, the attacker could pick up the signal and take responsive actions, such as switching C2 channel or pivoting to another system. It is also important to be aware of what happens behind the scenes of the tools used during the analysis of an event, as some of them might give away the status of the investigation without the analyst being aware of it. It is important to remember that if the Blue Team would have taken remediation actions to stop the attacker once detected then the Red Team would have been immediately alerted that the Blue Team is onto them. From an attacker perspective, it is important to keep monitoring OPSEC failures indicators in order to be able to quickly respond and increase their advantage on the defenders.

Interestingly the technology that an attacker can use to monitor their infrastructure for signs of Blue Team investigations is the same that the defenders use to

protect their systems. The concept of a Red Team monitoring system has been developed in open source projects such as RedElk (as it is described in section 6.2.1). However, it did not receive attention from the academic community yet. The last step for this research is to study which of these technologies can be used to detect OPSEC failures. Note that in the follow-up interview the Red Team did not mention any specific measure they take to detect OPSEC failures, however, they pointed out RedElk as a possible viable option (see section 6.2.1 for a description of RedElk)

Chapter 6

Red Team Infrastructure

The following chapter will present the result of brief research on the Red Team infrastructure. In order to answer to *RQ1*, it is necessary to understand what is the defence perimeter of the attacker, and subsequently, it will be possible to individuate technologies that can be used to detect IoCs that the Blue Team generates when touching such perimeter.

The results were obtained by web research, browsing through Red Team forums and online repositories. This way, we obtained a basic outline of the infrastructure necessary to support Red Team operations. The results were then validated with the Norhtwave Red Team and by inquiring with them about possible differences in the infrastructure they use.

6.1 Infrastructure

A typical Red Team assessment campaign can lasts weeks or months, and the probability of being detected increases over time. The Blue Team is continuously monitoring the system and actively seeking for threats within the network. The typical RT infrastructure takes into consideration all these elements. The following section will describe the core elements of the infrastructure sustaining Red Team operations.

Command and Control Server C2 is a system which enable the Red Team operator to control the implants on the victim machines. The malware establishes a com-

munication channel with the C2 server, and uses it to receive commands, download additional components or to exfiltrate information. There are different techniques and protocols used for the C2 channel. For long term Red Team operations is good practice adopting both a short haul and a long haul C2 server. The short haul C2 is typically used for ”every day” operations, based on how the malware is configured it contacts the short haul C2 periodically to see if the operator has issued any new command. On the other hand the long haul C2 server is a backup C2 used to maintain access on the exploited systems on the long term. The agent contacts the long haul C2 server very rarely and it is typically used only to pull new configurations in case the short haul C2 server has been detected and blacklisted. The two C2 differ also in the type of covert channel used, while the first might use any protocol from HTTP to SMTP, the long haul tries to use more stealthy ones such as DNS.

C2 covert channels The Red Team must maintain their OPSEC in order to perform a successful assessment. Therefore all the communications between the implants and the c2 must be hidden. Encrypting the data is not enough to avoid raising the suspicions of the Blue Team. If the BT detects suspicious activity, they can block it at the network and firewall level. A covert channel is used to communicate secretly and avoid being detected by hiding malicious traffic into legitimate traffic. It is possible to use different protocols to create covert channels, and the only limitation is the creativity of the operator, however there are some widely adopted ones such as HTTP, SMTP and DNS.

Redirectors If the Blue Team manages to detect the IP address of the C2 server, they can easily blacklist it, or even request the hosting provider to take down the server. In this scenario, the whole operation is compromised, and it is necessary to restart from zero. A common practice to protect the C2 server is to utilize redirectors. The idea is that of puting small machines easily replaceable if burned in front of the Red Team servers. Redirectors are web servers configured just to forward the requests based on some parameters. If the incoming connection does not match the parameters set for the implants, the user is redirected to a legitimate website, whereas if it matches the implant signature signature, it is redirected to the malicious server. An example of such signature is the use of a specific user-agent in the http header.

Payload and Phishing Server A core element of every Red Team assessment is the phishing campaign. A Phishing server is configured so that the emails sent are not flagged as spam by the email providers. Often couple with phishing server is the payload server. A typical scenario is a phishing email with included a link to the payload server to download the malware. This technique is used because email attachments are often scanned for the presence of malicious files. A payload server is typically used to delivery malware, however it might be used to hosts other files as well such as tools or configuration files.

Cracking machine The Red Team often obtains password hashes during their assessment that they have to crack. In this situation, a cracking machine is used. It consists of a workstation with a lot of GPU power. GPUs are used in this kind of machine because they are extremely efficient in computing parallel process, such as computing multiple hashes at the same time.

Domain The final element required to complete each Red Team infrastructure is one or more credible domains. Such domains should resemble legitimate domains and are chosen among the ones that are not already flagged as malicious; this way, the Red Team operation can remain stealth longer. Domain fronting is a technique that exploits CDN[1] to hide the attacker traffic to a specific website by cloaking it as a different domain.

6.2 Detection technologies

The elements presented in section 6.1 are the elements that can be investigated by the Blue Team. A technology that is able to detect Blue Team OPSEC failure should be able to monitor interactions with these elements.

Canary tokens Canary tokens are digital tripwires that can be embedded in different places. When someone interacts with them, they trigger an alert. These elements are pretty common even outside the cybersecurity realm, as a matter of fact, webbugs are trackers embedded in emails and used in marketing campaigns to

[1] Content Delivery Networks

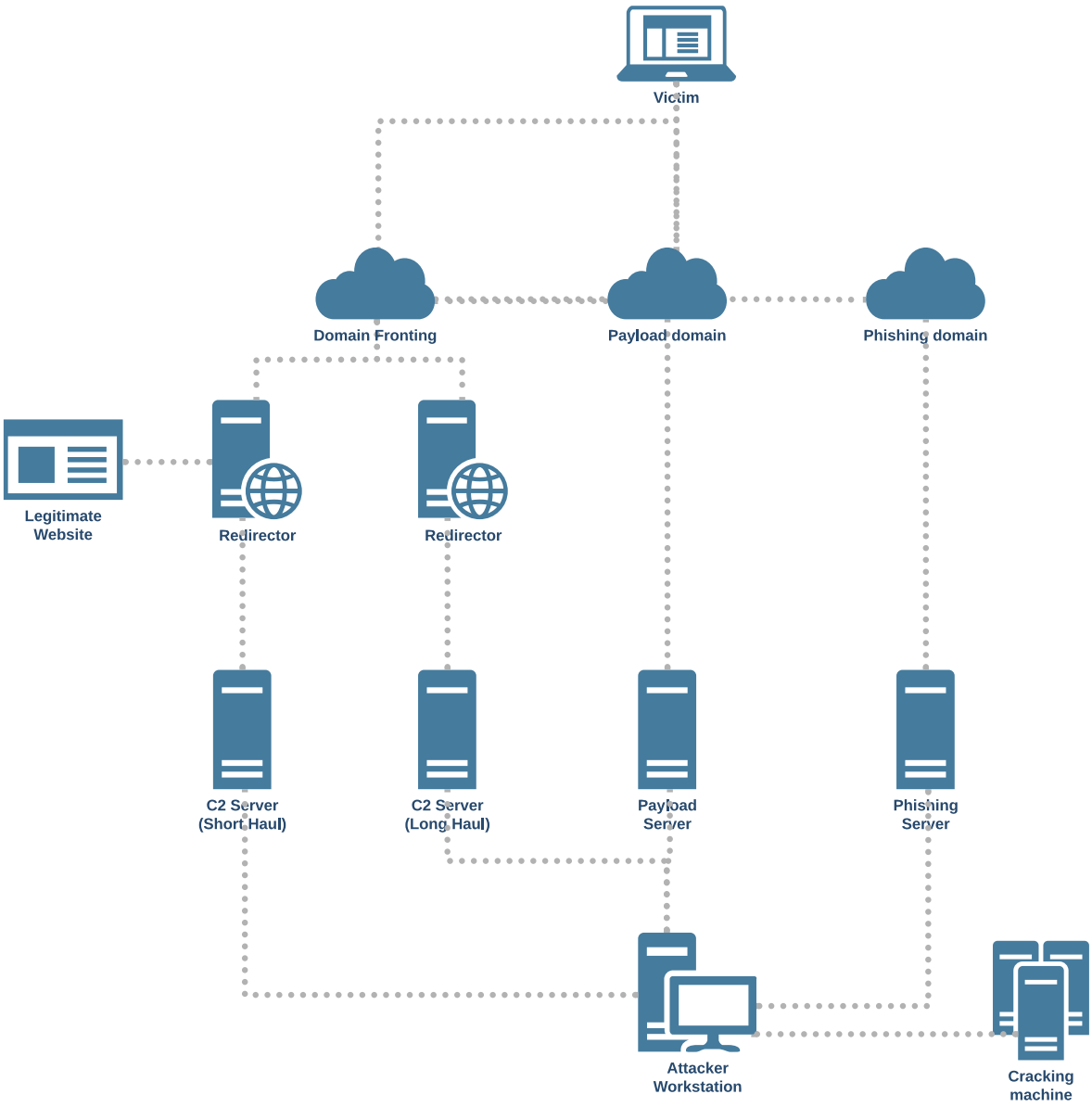


Figure 6.1: Interaction between the BT controlled variable and RT uncontrolled variable in the virtual environment

have a feedback on the number of users that have actually read the email. For example, it is possible to include in the email a link to an image hosted on a web server, when the user opens the email, the image is loaded, and the webserver generates an alert. Canary tokens exploit the same concept, with the difference that the can be embedded in any other object. It is possible to embed them in documents such as pdf or word documents triggering when they are accessed, in applications and triggers when they are reversed engineered, or in websites when they are cloned [88]. For instance, the tokens can be embedded in the Red Team's implant and generate ad alert if an analyst is reversing it.

Monitor redirectors Another promising way to detect if the Red Team is being investigated is to monitor the redirectors for failed connection attempts or suspicious connections. In a basic HTTP redirector this translate into collecting access logs and analysing them. This solution comes with its own set of challenges. Just like other monitoring systems IP based, the biggest challenge is to filter the white internet noise. However, being IP based detection so widely used and study that adapt one of the existing vendor tools to the Red Team usecase should be fairly trivial.

Monitoring Public services As observed in section 4.3.2, a possible OPSEC failure is the upload of malware to public services or public scans of suspicious websites. All these public services usually offer APIs that (premium) users can use to query their database. An attacker can easily set up a script that periodically queries such services for indicators that their malware or website has been scanned. Some of the services that can be monitored are: *VirustTotal* [89], *urlscan.io* [90], *any.run* [91],*Spamhaus* [92],*IBM X-force* [93].

Monitor unusual credentials Even though carefully inspecting the credentials received during a phishing campaign is not a detection technology, it is still a possible method the Red Team can use to be alerted about possible on-going SOC investigations. This method can be even handier in case the Red Team has performed the assessments against the same SOC before. Fake credentials are not changed with the same frequency as real credentials. They might even remain the same for a long time as there is no perceived threat in reusing bogus credentials. However, such credentials might be peculiar sometimes and easily stand out.

Monitor suspicious changes in system Changes in the system, such as the creation of new high privileged accounts, are clear indicators that an investigation is being performed. The attacker should monitor for such changes once he/she has established persistence on the system.

Targeted ads An interesting detection technique that has been proposed by a security researcher during DefCon 2018 [94] is to use targeted Google ads to spot Blue Team investigations. The idea behind his research is that analysts perform the investigation using a series of tools, first internal tools, then vendor tools and finally, public tools. As a last resort, a Google search is still the most valuable method to dig some more information when the other methods failed. This method involves embedding a specific, unique and not meaningless string in the malware code. Such a string should catch the eye of the investigator, for example, the name of a fictitious hacking group or an email address can be used. Moreover, in order to obtain clear-cut results, it should be a string that generates low search volume. The next step is to create a legitimate website containing that string. The last step is to create a google ad that is shown whenever someone searches for that particular string. This way it is possible to trigger an alert whenever someone googles for that string. There are some caveats with this technique. The main problem is that Google often changes the ad algorithm so it might not always so the detection might not always be reliable. Another problem is that register website and Google ads require the attacker to expose himself. Possibly compromising the OPSEC of his own operation.

6.2.1 RedElk

RedElk is an open-source project that includes most of the technologies described above. Red Elk is a SIEM for Red Teams which is used to support Red Team operations by tracking Blue Team investigation and generating alarms. The tool collects specifics IOCs generated by the Blue Team, then it aggregates the data, enriches the it and display it through the ELK stack[2]. It collects logs from the redirctors, and queries security service provider (such as Virustotal, Spamhaus, any.run) for indicators of the attack 6.2.

[2]"ELK" is the acronym for three open source projects: Elasticsearch, Logstash, and Kibana. Elasticsearch is a search and analytics engine. Logstash is data injestor, Kibana lets users visualize data

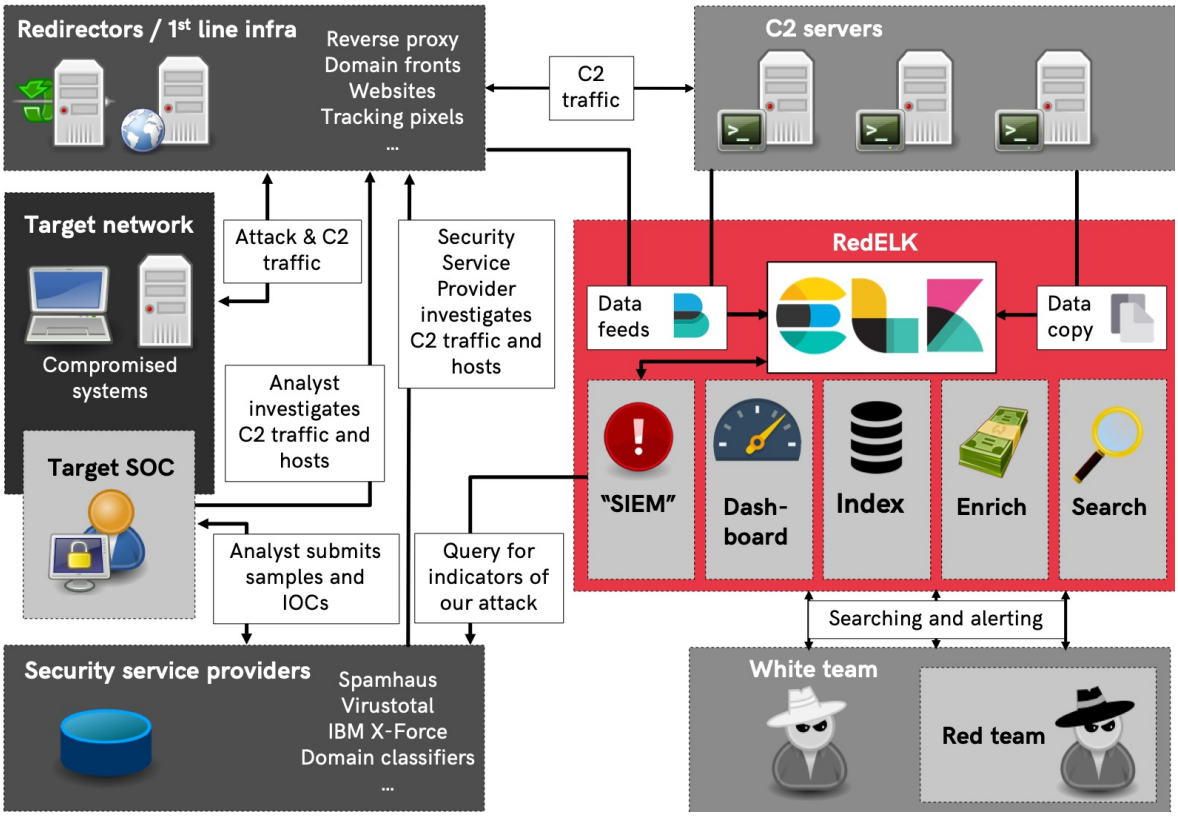


Figure 6.2: RedElk Overview

6.2.2 Conclusion

In conclusion, many different technologies can be used to spot Blue Team OPSEC failures. Even though all this technologies are very promising they are not well developed yet(except for RedElk). At the time of writing the best answer to *RQ1* is precisely to use RedElk. Nevertheless, if further developed the other technologies have the potential to be very beneficial as well. However, it is essential to note that implementing all these measures requires a lot of extra work for an attacker, mostly fine-tuning time to obtain more reliable results. Adopting measures to detect OPSEC failures is, therefore, an approach that makes sense for highly motivated attackers or state-sponsored hacking groups with big budgets and resources to invest (as in section 2.1.4). These actors are highly motivated in protecting their investment.

Chapter 7

Conclusion

7.1 Answering Research Questions

Research Question 2

What are the most common SOC analysts OPSEC failures?

During ”Phase I” of this research project, a mixed Cognitive task analysis approach was used to access analyst tacit knowledge about their investigative process. Together with the analysts, we examined some hypothetical cyber-attack scenarios from the perspective of ”unware” SOC analysts. In doing so, we urged the interviewee into highlighting and critiquing the mistakes such hypothetical analyst could make. We categorised the errors they listed into three categories: actions that can *compromise the investigations*, actions that result in a direct *risk for the analyst*, and finally actions that can cause *indirect damage to the customer*. However, due to the fact that we limited the research to a single SOC (Northwave's SOC) it is not possible to state that the highlighted OPSEC failures are indeed the ”most” common. Nonetheless, we can classify them as ”likely to happen”, especially considering the highly stressful environment where the SOC analysts operate.

Research Question 3

To what extent the Red Team is aware of the Blue Team investigation?

The first step to answer this question was to observe possible anomalous behaviours of the attackers during the wargame. Where anomalous refers to the canonical steps of an attack defined by the Unified Cyber Kill Chain. No major change of strategies have been observed for the teams. However, one of the teams tried to exploit the presence of the SOC by making a phone call to one of the members of the SOC and trying a social engineering attack. This reveals that the Blue Team itself is viewed as an opportunity as much as a rival. On the other hand this does not indicate that the team was aware of the status of the Blue Team investigation.

Based on the answers during the follow-up interviews, there are indications that the Red Team was aware that they were being investigated. However, at the beginning of the wargame, they were informed that no active countermeasure would have been taken to stop them (due to time constraints). This fact made them perceive the Blue Team as less of a treat. Hence, giving the impression of not being aware of their investigation. In general, the teams did not adopt a monitoring system for their infrastructure. Therefore they are not aware of the status of the investigation until some countermeasure is taken to stop them.

Research Question 1

How the Red Team can detect Blue Team OPSEC failures?

We discussed in chapter 6 about the possible technologies that can be used to detect OPSEC failures. The Red Team either monitor their own infrastructure, monitor online services or inspecting carefully the data provided by the victims. There can be many more technologies that can be used. Those the technologies we presented were specifically tailored to pickup OPSEC failures of SOC analysts. After discussing the possible technologies(6), we concluded that the best option for the Red Team is to adopt free solution such as RedElk. This way, without investing a significant amount of time in acquiring and tuning tools, they can have some more insights into Blue Team operations.

Additional Results

Researching a previously almost unexplored field revealed to be a rather challenging task. Additionally, to the research questions previously mentioned, this project had

the ambitious goal of creating the groundwork for future research in the field of attack-defender interplay. We had to develop most of the tools required to answered the research questions; each one of these elements is a unique contribution to the knowledge base for future academics.

CTA Methodology

The first issue we encountered was how to access tacit knowledge of the professionals involved in the research (i.e the SOC analysts). We observed that the analysts processed high amounts of alarms every day and that the more experienced analysts developed an investigative process that is highly effective. We aimed at identifying flaws in such a process that the Red Team could exploit. However, we observed that the average analyst goes through each step so fast that describing at which point a mistake could be made was a challenge. Therefore we developed and validated a Cognitive Task Analysis method to elicit knowledge from the analysts. The process is a mix of semi-structured interviews, hypothetical scenario analysis, and elicitation by critiquing.

Reusable Wargame Infrastructure

In the effort of validating the findings of the interviews, we developed a reusable and realistic wargame environment. Distinctive aspects of the environment are that it does not require underlying physical infrastructure, and it is modular, therefore, can be used in many different situations. The other unique feature is that it includes a built-in Security Operation Center. This feature is rarely included even in the more complex projects. Additionally, the low cost allows it to be the perfect starting point to support future researchers. We believe that providing an easy way to set up a wargame would stimulate researchers to explore this field even more. At the time of writing, most of the research on attackers and defenders strategies happens during big cyber wargames, which are rare and upon which the researcher has no control.

Red Teaming

On the Red Team side of the research, we defined what means Red Teaming, explained in detailed what are its origins and the difference with Pentesting. We also

explored different Red Teaming standard adopted worldwide and highlighted common presenting a unified definition of the process. We did this analysis on Red Teaming because we believed that it is not possible to research the dynamics between two actors if one has not fully understood the essence of each one of them. Furthermore, Red Teaming was understudied and received little academic attention over the years.

Blue Team Variable

On the Blue Team side, we experimented a new approach where their actions were modelled as a variable that could be used to stimulate certain actions in the adversary. As far as we are aware of this was the first attempt of using this method in this context. We modelled this variable as a flowchart rather then of a static list of actions in order to reflect the dynamic nature of the subject.

Active Partecipation of the Researcher

The final contribution was demonstrating benefits (and the disadvantages) when the researcher participates proactively rather than being a passive observer in this kind of research. The benefits were evident during the interviews phase as the experience of the researcher as Tier1 analysts played a role in the design of the hypothetical scenarios. The high control over the action of the Blue Team during the experiment will allow future researchers to easily reproduce the results of our research. Such level of control would not have been possible without actively involving the researcher in the experiment. However, there were some drawbacks during the wargame due to lack of workforce. Including other researchers in the experiment in order to be able to manage more aspects of the wargame would require to train them as SOC analysts first, and this task is highly time-consuming.

7.2 Future Work

This research aimed at creating the groundwork for future research on attacker and defender interplay. Being this a mostly unexplored field, there is much work left to do to improve the field. There are many different directions future work can take

starting from where this research stopped. This section will present some of the most relevant directions that are possible to take from this point. However, the tools developed in this research can be further refined and used for any kind of research.

The whole research was tailored on SOC analysts; we considered their problems, their investigative process and the technologies they used to support their operations. However, there are many more sub-teams in the Blue Teams that are worth of being further studied. For instance, this research can be easily extended by studying Threat Hunters and how their operations can affect the Red Team.

Talking about the Blue Team, it would be interesting to allow a CERT team to investigate the environment after a wargame. Due to the fact that the environment is both easy to deploy but also to clone, an interesting research direction would be to use it to evaluate the ability of different CERT teams. Using the same cloned testbed for the evaluation would allow them to obtain comparable results.

This research focused on the interplay between attacker and defenders; a logical next step is to study the influence of the Red Team on the Blue Team operations. In this way, we are moving one step closer to obtain a holistic view of the relationship between the attacker and defenders.

After having improved the environment according to the indication in section 5.2.3, it would be possible to repeat the experiment. However, it is possible to change configurations such as software installed on the workstations, However, trying different network configurations and observing how different network topologies affect the strategies of the adversaries.

A final research question. Would it be possible to automate the copy of existing real infrastructure and then redeploy a virtual version of it?. A positive answer to this research question would have a number of disruptive implications both from an academic perspective and from an industry perspective. It would be possible to perform a complete Red Team assessment in a safe environment without the risk of causing damage to the tested organisation.

Bibliography

[1] M. Mateski, “Red teaming and alternative analysis.” [Online]. Available: https://redteamjournal.com/red-teaming-and-alternative-analysis

[2] H. H. Friedman, “Cognitive biases that interfere with critical thinking and scientific reasoning: A course module,” *Available at SSRN 2958800*, 2017.

[3] I. L. Janis, “Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.” 1972.

[4] P. Hart, “Irving l. janis’ victims of groupthink,” *Political Psychology*, vol. 12, p. 247, 06 1991.

[5] J. Klayman, “Varieties of confirmation bias,” *Psychology of learning and motivation*, vol. 32, pp. 385–418, 1995.

[6] V. M. Tripwire Guest AuthorsFeb 27, “Red teaming: How to run effective cyber-drills?” Feb 2020. [Online]. Available: https://www.tripwire.com/state-of-security/vulnerability-management/red-teaming-effective-cyber-drills/

[7] S. NIST, “800-53, rev. 4,“security and privacy controls for federal information systems,” national institute of standards and technology, april 2013.”

[8] “The penetration testing execution standard,” 2014, PTEST. [Online]. Available: http://www.pentest-standard.org/index.php?title=Main_Page&oldid=950

[9] K. Scarfone, M. Souppaya, A. Cody, and A. Orebaugh, “Nist special publication 800-115: Technical guide to information security testing and assessment,” *Maryland: National Institute of Standards and Technology*, 2008.

[10] P. Herzog, “Osstmm 3-the open source security testing methodology manual: Contemporary security testing and analysis,” *ISECOM-Institute for Security and Open Methodologies*, 2010.

[11] P. Brangetto, E. Çalişkan, and H. Rõigas, “Cyber red teaming,” *NATO Cooperative Cyber Defence Centre of Excellence CCDCOE*, 2015.

[12] H. Dalziel, *Next Generation Red Teaming*. Syngress, 2015.

[13] "Red Team: Adversarial Attack Simulation Exercise Guidelines for the Financial Industry in Singapore," Association of Banks in Singapore (ABS), Tech. Rep., 2018.

[14] Wikipedia contributors, "Modus operandi — Wikipedia, the free encyclopedia," 2020, [Online; accessed 8-April-2020]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Modus_operandi&oldid=949461689

[15] L. Ablon, *Data Thieves: The Motivations of Cyber Threat Actors and Their Use and Monetization of Stolen Data*. RAND Corporation, 2018. [Online]. Available: http://dx.doi.org/10.7249/CT490

[16] H. C. Robinson,Neil, Gribbon,Luke, *Cyber-security threat characterisation: A rapid comparative analysis*. RAND Corporation, 2013. [Online]. Available: https://www.rand.org/pubs/research_reports/RR235.html

[17] M. Kenney, "Cyber-terrorism in a post-stuxnet world," *Orbis*, vol. 59, no. 1, p. 111–128, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.orbis.2014.11.009

[18] E. M. Hutchins, M. J. Cloppert, and R. M. Amin, "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," *Leading Issues in Information Warfare & Security Research*, vol. 1, no. 1, p. 80, 2011.

[19] R. Stolte, "Reactive or proactive? making the case for new kill chains," 2018. [Online]. Available: https://www.darkreading.com/attacks-breaches/reactive-or-proactive-making-the-case-for-new-kill-chains-/a/d-id/1332200

[20] G. Engel, "Deconstructing the cyber kill chain," Dec 2014. [Online]. Available: https://www.darkreading.com/attacks-breaches/deconstructing-the-cyber-kill-chain/a/d-id/1317542

[21] P. Pols, "The unified kill chain: Designing a unified kill chain for analyzing, comparing and defending against cyber attacks," *Cyber Security Academy*, 2017.

[22] K. Oosthoek and C. Doerr, "Sok: Att&ck techniques and trends in windows malware," in *International Conference on Security and Privacy in Communication Systems*. Springer, 2019, pp. 406–425.

[23] B. E. Strom, J. A. Battaglia, M. S. Kemmerer, W. Kupersanin, D. P. Miller, C. Wampler, S. M. Whitley, and R. D. Wolf, "Finding cyber threats with att&ck-based analytics," *The MITRE Corporation, Tech. Rep.*, 2017.

[24] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, "Mitre att&ck: Design and philosophy," *MITRE Product MP*, pp. 18–0944, 2018.

[25] A. Zakrzewski, T. Tang, G. Appell, R. Fages, A. Hardie, N. Hildebrandt, M. Kahlich, M. Mende, F. Muxi, and A. Xavier, "Global Wealth 2019: Reigniting Radical Growth," Tech. Rep. [Online]. Available: https://image-src.bcg.com/Images/BCG-Reigniting-Radical-Growth-June-2019_tcm9-222638.pdf

[26] J. Creasey, "A guide for running an effective Penetration Testing programme," Association of Banks in Singapore (CREST), Tech. Rep., 2017. [Online]. Available: https://www.crest-approved.org/wp-content/uploads/CREST-Penetration-Testing-Guide.pdf

[27] "Threat intelligence-based ethical red teaming," European Central Bank (BCE), Tech. Rep., 2018. [Online]. Available: https://www.ecb.europa.eu/pub/pdf/other/ecb.tiber_eu_framework.en.pdf

[28] "Intelligence-led cyber attack simulation testing," Tech. Rep., 2018. [Online]. Available: https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2016/20161221e1.pdf

[29] "Financial entities ethical red teaming," Saudi Arabian Monetary Authority (SAMA), Tech. Rep., 2019. [Online]. Available: http://www.sama.gov.sa/en-US/Laws/BankingRules/FinancialEntitiesEthicalRedTeamingFramework.pdf

[30] J. Prenio, J. Yong, and R. Kleijmeer, "Varying shades of red: how red team testing frameworks can enhance the cyber resilience of financial institutions," 2019.

[31] M. Khalili, "Monitoring and improving managed security services inside a security operation center," Ph.D. dissertation, Concordia University, 2015.

[32] A. Michail, "Security operations centers: A business perspective," Master's thesis, 2015.

[33] C. Zhong, J. Yen, P. Liu, and R. F. Erbacher, "Automate cybersecurity data triage by leveraging human analysts' cognitive process," in *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)*. IEEE, 2016, pp. 357–363.

[34] M. A. Champion, P. Rajivan, N. J. Cooke, and S. Jariwala, "Team-based cyber defense analysis," in *2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*. IEEE, 2012, pp. 218–221.

[35] S. C. Sundaramurthy, J. Case, T. Truong, L. Zomlot, and M. Hoffmann, "A tale of three security operation centers," in *Proceedings of the 2014 ACM workshop on security information workers*, 2014, pp. 43–50.

[36] c. sanders and s. rand, "creative choices: developing a theory of divergence, convergence, and intuition in security analysts."

[37] "Operations security," 2018. [Online]. Available: https://aglearn.usda.gov/customcontent/APHIS/APHIS-OPSEC/OPSsummary.htm

[38] K. A. Seger, *Utility security: the new paradigm*. PennWell Books, 2003.

[39] J. T. Force and T. Initiative, "Security and privacy controls for federal information systems and organizations," *NIST Special Publication*, vol. 800, no. 53, pp. 8–13, 2013.

[40] L. Cholvy, "Information evaluation in fusion: Formalization of informal recommendations," in *Modern Information Processing*. Elsevier, 2006, pp. 245–254. [Online]. Available: https://doi.org/10.1016/b978-044452075-3/50021-2

[41] K. Scarfone, W. Jansen, and M. Tracy, "Guide to general server security," *NIST Special Publication*, vol. 800, no. s 123, 2008.

[42] D. Miessler and D. Miessler, "Opsec is obscurity, and opsec increases security," Dec 2019. [Online]. Available: https://danielmiessler.com/blog/opsec-is-obscurity-and-opsec-increases-security/

[43] W. T. Johnsen, *The Principles of War in the 21st Century: Strategic Considerations*. DIANE Publishing, 1995.

[44] D. A. Wheeler and G. N. Larsen, "Techniques for cyber attack attribution," INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA, Tech. Rep., 2003.

[45] J. Hunker, B. Hutchinson, and J. Margulies, "Role and challenges for sufficient cyber-attack attribution," *Institute for Information Infrastructure Protection*, pp. 5–10, 2008.

[46] T. Rid and B. Buchanan, "Attributing cyber attacks," *Journal of Strategic Studies*, vol. 38, no. 1-2, pp. 4–37, 2015.

[47] D. D. Clark and S. Landau, "Untangling attribution," *Harv. Nat'l Sec. J.*, vol. 2, p. 323, 2011.

[48] M. Fabro, L. P. Inc, and V. Maio, "Using operational security(opsec) to support acyber security culture in control systems environments version 1.0," *INL Critical InfrastructureProtection Center*, vol. 22, September 2007.

[49] J. C. Dressler, C. Bronk, and D. S. Wallach, "Exploiting military opsec through open-source vulnerabilities," in *MILCOM 2015-2015 IEEE Military Communications Conference*. IEEE, 2015, pp. 450–458.

[50] 2020. [Online]. Available: https://github.com/outflanknl/RedELK

[51] S. Xu, "Cybersecurity dynamics: A foundation for the science of cybersecurity," in *Proactive and Dynamic Network Defense*. Springer, 2019, pp. 1–31.

[52] F. He, S. Chandrasekar, N. S. V. Rao, and C. Y. T. Ma, "Effects of interdependencies on game-theoretic defense of cyber-physical infrastructures," *2019 22th International Conference on Information Fusion (FUSION)*, pp. 1–8, 2019.

[53] R. Luh, M. Temper, S. Tjoa, S. Schrittwieser, and H. Janicke, "Penquest: a gamified attacker/defender meta model for cyber security assessment and education," *Journal of Computer Virology and Hacking Techniques*, vol. 16, pp. 19 – 61, 2019.

[54] J. M. Haney and C. L. Paul, "Toward integrated tactical operations for red/blue cyber defense teams."

[55] M. Branlat, "Challenges to adversarial interplay under high uncertainty: staged-world study of a cyber security event," Ph.D. dissertation, The Ohio State University, 2011.

[56] J. Vykopal, M. Vizváry, R. Oslejsek, P. Celeda, and D. Tovarnak, "Lessons learned from complex hands-on defence exercises in a cyber range," in *2017 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2017, pp. 1–8.

[57] M. G. VISKY, "cyber-physical battlefield for cyber exercises," in *5th interdisciPlinary cyber research conference 2019*, 2019, p. 10.

[58] K. N. Lovell, "cyber game to cyber exercise: a new methodology for cybersecurity simulations," in *5th interdisciPlinary cyber research conference 2019*, 2019, p. 13.

[59] Wikipedia contributors, "Computer ethics — Wikipedia, the free encyclopedia," 2020, [Online; accessed 4-October-2020]. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Computer_ethics&oldid=981595660

[60] R. E. Pike, "The "ethics" of teaching ethical hacking," *Journal of International Technology and Information Management*, vol. 22, no. 4, p. 4, 2013.

[61] N. Radziwill, J. Romano, D. Shorter, and M. Benton, "The ethics of hacking: Should it be taught?" *arXiv preprint arXiv:1512.02707*, 2015.

[62] R. D. Hartley, "Ethical hacking pedagogy: an analysis and overview of teaching students to hack," *Journal of International Technology and Information Management*, vol. 24, no. 4, p. 6, 2015.

[63] P. Johannesson and E. Perjons, *An introduction to design science*. Springer, 2014.

[64] "Northwave. intelligent security operations," Jul 2020. [Online]. Available: https://northwave-security.com/

[65] B. Crandall, G. Klein, G. A. Klein, R. R. Hoffman *et al.*, *Working minds: A practitioner's guide to cognitive task analysis*. Mit Press, 2006.

[66] K. Gallo, "Understanding professional jargons literature review," 10 2016.

[67] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, and K. Hoagwood, "Purposeful sampling for qualitative data collection and analysis in mixed method implementation research," *Administration and policy in mental health and mental health services research*, vol. 42, no. 5, pp. 533–544, 2015.

[68] J. E. Miller, E. S. Patterson, and D. D. Woods, "Elicitation by critiquing as a cognitive task analysis methodology," *Cognition, Technology & Work*, vol. 8, no. 2, pp. 90–102, 2006.

[69] K. E. Heckman, M. J. Walsh, F. J. Stech, T. A. O'boyle, S. R. DiCato, and A. F. Herber, "Active cyber defense with denial and deception: A cyber-wargame experiment," *computers & security*, vol. 37, pp. 72–77, 2013.

[70] D. B. Fox, C. D. McCollum, E. I. Arnoth, and D. J. Mak, "Cyber wargaming: Framework for enhancing cyber wargaming with realistic business context," MITRE CORP MCLEAN VA MCLEAN, Tech. Rep., 2018.

[71] A. D'Amico, K. Whitley, D. Tesone, B. O'Brien, and E. Roth, "Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts," in *Proceedings of the human factors and ergonomics society*

annual meeting, vol. 49, no. 3. SAGE Publications Sage CA: Los Angeles, CA, 2005, pp. 229–233.

[72] F. B. Kokulu, A. Soneji, T. Bao, Y. Shoshitaishvili, Z. Zhao, A. Doupé, and G.-J. Ahn, "Matched and mismatched socs: A qualitative study on security operations center issues," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 1955–1970.

[73] M. Granåsen and D. Andersson, "Measuring team effectiveness in cyber-defense exercises: a cross-disciplinary case study," *Cognition, Technology & Work*, vol. 18, no. 1, pp. 121–143, 2016.

[74] R. S. S. Kumar, A. Wicker, and M. Swann, "Practical machine learning for cloud intrusion detection: challenges and the way forward," in *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security*, 2017, pp. 81–90.

[75] J. Wei and G. Salvendy, "The cognitive task analysis methods for job and task design: Review and reappraisal," *Behaviour & Information Technology*, vol. 23, no. 4, pp. 273–299, 2004.

[76] L. Argote and P. Ingram, "Knowledge transfer: A basis for competitive advantage in firms," *Organizational behavior and human decision processes*, vol. 82, no. 1, pp. 150–169, 2000.

[77] S. Harries, "3 - concepts, codes and meanings: bridging knowledge and records," in *Records Management and Knowledge Mobilisation*, ser. Chandos Information Professional Series, S. Harries, Ed. Chandos Publishing, 2012, pp. 49 – 66. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9781843346531500032

[78] P. R. Gamble and J. Blackwell, *Knowledge management: A state of the art guide*. Kogan Page Publishers, 2001.

[79] W. Swap, D. Leonard, M. Shields, and L. Abrams, "Using mentoring and storytelling to transfer knowledge in the workplace," *Journal of management information systems*, vol. 18, no. 1, pp. 95–114, 2001.

[80] L. Militello and R. Hutton, "Applied cognitive task analysis (acta): A practitioner's toolkit for understanding cognitive task demands," *Ergonomics*, vol. 41, pp. 1618–41, 12 1998.

[81] [Online]. Available: https://ccdcoe.org/exercises/locked-shields/

[82] T. W. Edgar and T. R. Rice, “Experiment as a service,” in *2017 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, 2017, pp. 1–6.

[83] J. Vykopal, R. Ošlejšek, P. Čeleda, M. Vizvary, and D. Tovarňák, “Kypo cyber range: Design and use cases,” 2017.

[84] M. M. Yamin, B. Katt, and V. Gkioulos, “Cyber ranges and security testbeds: Scenarios, functions, tools and architecture,” *Computers & Security*, vol. 88, p. 101636, 2020.

[85] M. Wittig, A. Wittig, and B. Whaley, *Amazon web services in action*. Manning, 2016.

[86] E. G. Gerosa, “Blueteamlabs sentinel-attack.” [Online]. Available: https://github.com/BlueTeamLabs/sentinel-attack

[87] “Sysmon - windows sysinternals.” [Online]. Available: https://docs.microsoft.com/en-us/sysinternals/downloads/sysmon

[88] “Canary tokens.” [Online]. Available: https://canarytokens.org/

[89] VirusTotal. [Online]. Available: https://www.virustotal.com

[90] Urlscan.io, “urlscan.io.” [Online]. Available: https://urlscan.io/

[91] any.run, “any.run.” [Online]. Available: https://any.run/

[92] “The spamhaus project.” [Online]. Available: https://www.spamhaus.org/

[93] “Ibm x-force,” Oct 2020. [Online]. Available: https://www.ibm.com/nl-en/marketplace/ibm-xforce-exchange

[94] “Detecting blue team research through targeted ads,” Presentation at DefCon Las Vegas 2018, 2018.

Appendix A

Interviews

A.1 Semi-structured Interviews

The SOC analysts were interviewed using the semi-structured interview format. A series of open questions were prepared and asked to the participants, in some cases if the interviewee brought up an interesting topic additional questions were asked which were not prepared. There were multiple goals in this first part of the interview. The first one was to profile the participants and their background. The second one was inquire with the participants about their workflow and verify if any major difference existed among different analysts. The last goal was to let analyst talk and slowly think about a process that most of the time they perform very fast and often become an automatism. The last goal exists to prepare the ground for the second part of the interview, which is the hypothetical scenarios analysis. In this way the participant would start the analysis with the process they follow clear in their minds and most importantly they had the time to think about the reasons why they made certain choices. This approach was not documented in the other publications which were analyzed, and can therefore be considered an important contribution to future research in the field.

Following the list of questions which were asked to the interviewee.

Question 1 What is your role whithin the SOC? Briefly explain what it is and what are your main duties.

Question 2 How much experience do you have as SOC analyst? Can you define a threshold of experience up to a SOC analyst is a novice, and beyond that is an expert?

Question 3 Which kind of SOC/analyst related training do you have? Do you have any certification?

Question 4 Can you give an high-level description of your investigative process?

Question 5 Which measures do you take to preserve your OPSEC? And where did you learn about this practices?

Question 6 The scheme in figure 4.1 was presented to the candidates, and a brief explanation was given.

- When investigate an incident do you go through each one of this stages?
- Which steps are the most automated?
- During an investigation which of this stages takes you most time?
- At which step is more error prone in your opinion? and why?

A.2 Hypothetical scenarios

During the interviews the interviewee were presented with five different hypotetical scenarios, and they were asked to describe the analysis they would make in that situation. Each interviewee was then specifically asked to describe which mistakes an analyst could make when investigating the very same scenario. It is important to specify that the scenarios were not designed with the goal of evaluate the analyst's ability in performing the investigation. Whereas, the goal was to give them sufficient background and information to stimulate their memory and imagination and get good insights in their investigative proecess. Such choice is motivated by the fact that the intention of the interviews was not to evaluate the capabilities of the analysts, the

goal is instead of capture their knowledge. This scenario-based approach is typical Cognitive Task Analysis (CTA) techniques.

The scenario based analysis was structured as follow. The participants were asked to picture themselves in front of their SOC console. They were told that they could assume to have access to any tool or information they would normally have during their regular job (this step was eased by the the preliminary questions they answered during the first part of the interview). Then for each scenario they were asked to ”think out loud” all the steps they would follow to analyzed such scenario, and also to mention the indicators of malicious intent they would look for.

During the second part of the scenario based analysis the participant were asked to perform the very same analysis once again. However, this time they were asked to imagine to be a ”bad” SOC analysts, who is not aware of any OPSEC practice. They were, therefore, asked to emphasized and mentioned every possible mistake an analyst could make. At this stage it was very important that the participants could brainstorm an be creative with their answers.

All the scenarios were designed to follow the actions of the attackers from the moment they get the initial foothold in the network to

All the scenarios were designed to focus on the actions an attacker could do from getting the inital foothold to get the initial foothold in the network, or the steps that follow right after the initial foothold is established.

The scenarios were design to follow all the steps an attacker that tries to get into the system. However, they were mostly focused on the early steps when the attacker tries to get the initial foothold and to establish persistence in the system. The reason for that is that this is the most common situation a SOC analysts analyse. The subsequent steps of the kill chain are more often handled by threat hunters or other security teams.

The general structure of the interview should follow the steps of an attack along the cyber kill chain. For this reason figure A.1 maps each scenario to a different stage of the kill chain.

Following are described the five scenarios used in the interviews:

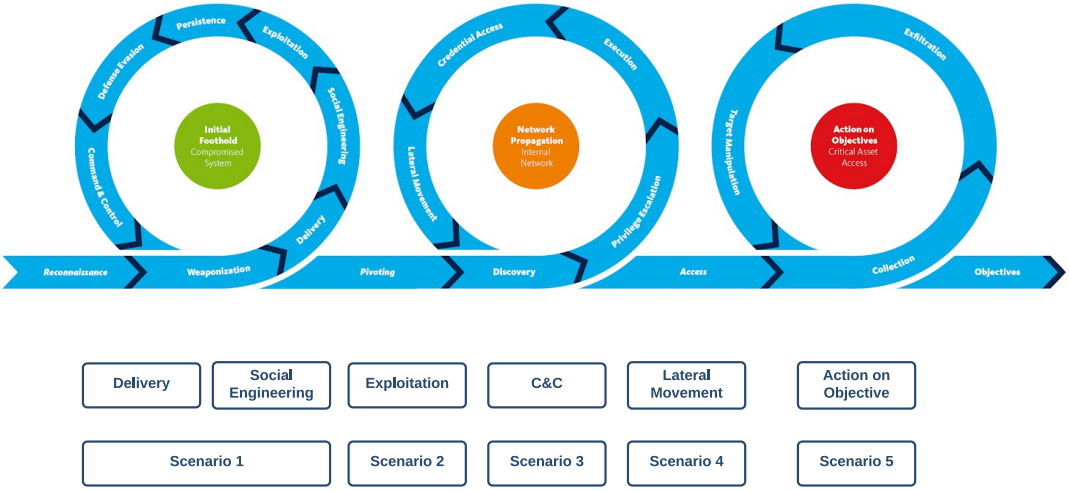


Figure A.1: Mapping of scenarios to Unified Cyber Kill Chain

A.2.1 Scenario 1

The first scenario presented to the SOC analysts was a phishing email. This was the more complex scenario amongst the various which were used during the interviews. It is composed of two steps, the first one being the analysis of the phishing email itself, and the second one being the analysis of phishing landing page. A suspicious email was reported to the Security Operation Center by one of the customers for further investigation. The email was written in good English, and there were several links.

The second step was to analyze the phishing website one of the link in the email was pointing to. Some of the hidden indicators were for instance the website name.

A.2.2 Scenario 2

The third scenario is fileless attack. In this scenario the SOC have been notified that a malicious document was received. Such document was a word document which contained indications that the user should enable macros in order to properly visualize the content.

Figure A.2: Scenario 1 - phishing email

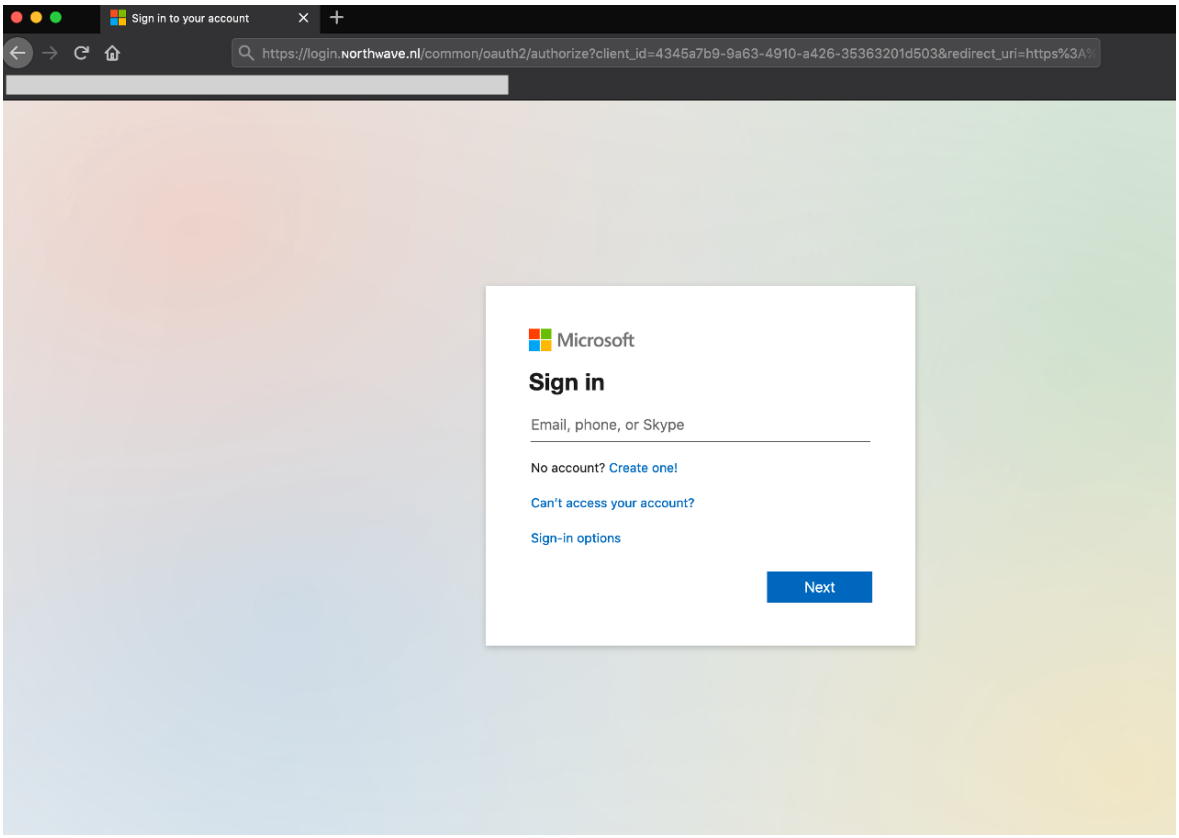


Figure A.3: Scenario 1 - phishing website

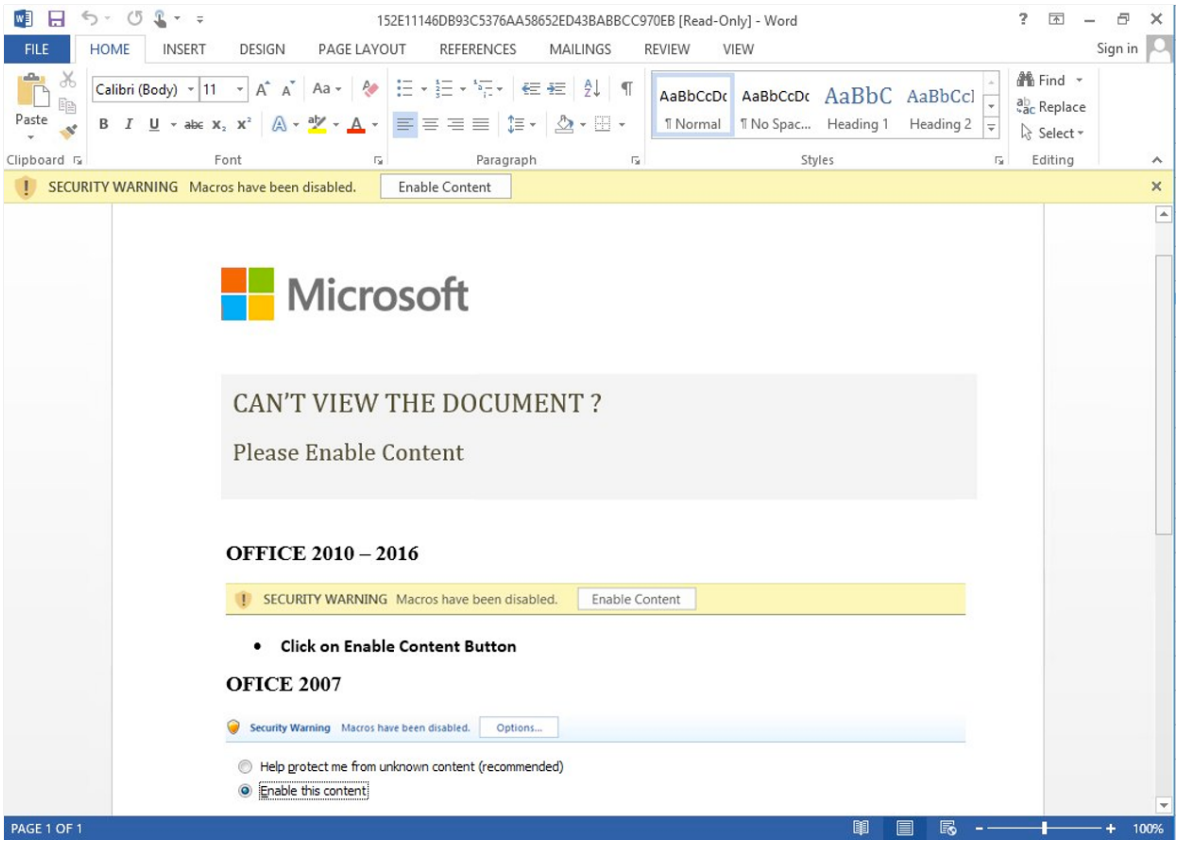
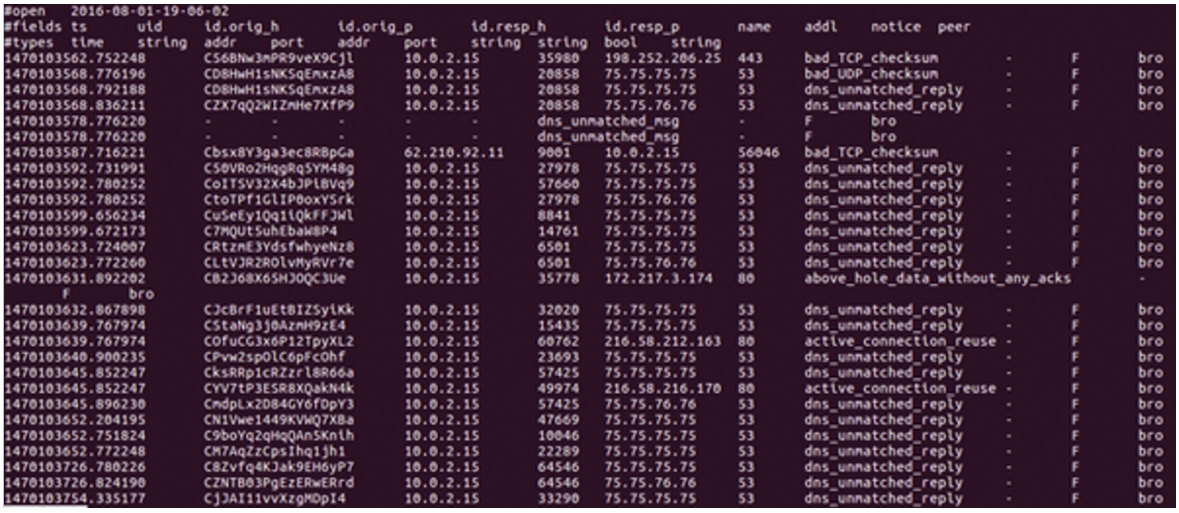


Figure A.4: Scenario 2 - fileless attack



```
#open   2016-08-01-19-06-02
#fields ts      uid     id.orig_h       id.orig_p       id.resp_h       id.resp_p       name    addl    notice  peer
#types  time    string  addr    port    addr    port    string  string  bool    string
1470103562.752248       CS6BNw3mPR9veX9Cjl      10.0.2.15       35980   198.252.206.25  443     bad_TCP_checksum        -       F       bro
1470103568.776196       CD8HwH1sNKSqEmxzA8      10.0.2.15       20858   75.75.75.75     53      bad_UDP_checksum        -       F       bro
1470103568.792188       CD8HwH1sNKSqEmxzA8      10.0.2.15       20858   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103568.836211       CZX7qQ2WIZmHe7XfP9      10.0.2.15       20858   75.75.76.76     53      dns_unmatched_reply     -       F       bro
1470103578.776220       -       -       -       -       dns_unmatched_msg       -       F       bro
1470103578.776220       -       -       -       -       dns_unmatched_msg       -       F       bro
1470103587.716221       Cbsx8Y3ga3ec8RBpGa      62.210.92.11    9001    10.0.2.15       56046   bad_TCP_checksum        -       F       bro
1470103592.731991       CS0VRo2HqgRqSYM48g      10.0.2.15       27978   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103592.780252       CoITSV32X4bJPl8Vq9      10.0.2.15       57660   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103592.780252       CtoTPf1GlIP0oxYSrk      10.0.2.15       27978   75.75.76.76     53      dns_unmatched_reply     -       F       bro
1470103599.656234       CuSeEy1Qq1iQkFFJWl      10.0.2.15       8841    75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103599.672173       C7MQUtSuhEbaW8P4        10.0.2.15       14761   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103623.724007       CRtznE3YdsfwhyeNz8      10.0.2.15       6501    75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103623.772260       CLtVJR2ROlvMyRVr7e      10.0.2.15       6501    75.75.76.76     53      dns_unmatched_reply     -       F       bro
1470103631.892202       CB2J68X65HJOQC3Ue       10.0.2.15       35778   172.217.3.174   80      above_hole_data_without_any_acks        -
        F       bro
1470103632.867898       CJcBrF1uEtBIZSyiKk      10.0.2.15       32020   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103639.767974       CStaNg3j0AzmH9zE4       10.0.2.15       15435   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103639.767974       COfuCG3x6P12TpyXL2      10.0.2.15       60762   216.58.212.163  80      active_connection_reuse -       F       bro
1470103640.900235       CPvw2spOlC6pFcOhf       10.0.2.15       23693   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103645.852247       CksRRp1cRZzrl8R66a      10.0.2.15       57425   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103645.852247       CYV7tP3ESR8XQakN4k      10.0.2.15       49974   216.58.216.170  80      active_connection_reuse -       F       bro
1470103645.896230       CmdpLx2D84GY6fDpY3      10.0.2.15       57425   75.75.76.76     53      dns_unmatched_reply     -       F       bro
1470103652.204195       CN1Vwe1449KVWQ7XBa      10.0.2.15       47669   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103652.751824       C9boYq2qHqQAnSKnih      10.0.2.15       10046   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103652.772248       CM7AqZzCpsIhq1jh1       10.0.2.15       22289   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103726.780226       C8Zvfq4KJak9EH6yP7      10.0.2.15       64546   75.75.75.75     53      dns_unmatched_reply     -       F       bro
1470103726.824190       CZNTB03PgEzERwERrd      10.0.2.15       64546   75.75.76.76     53      dns_unmatched_reply     -       F       bro
1470103754.335177       CjJAI11vvXzgMDpI4       10.0.2.15       33290   75.75.75.75     53      dns_unmatched_reply     -       F       bro
```

Figure A.5: scenario 3 - suspicious network traffic

A.2.3 Scenario 3

In the fourth scenario the analyst was asked to analyze network traffic logs. At this stage the assumption was that the attacker had established a command&control channel, and the analyst should be looking for suspicious network activities.

A.2.4 Scenario 4

In the second scenario the analyst had to analyze a password spraying attack. The participant was presented a picture of the interface of a SIEM with multiple failed login attempts.

A.2.5 Scenario 5

In the last scenario the interviewee could analyze the behaviour of a process. Differently from the previous scenarios the goal was to observer how the process of the analyst would change when investigating an attacker who was at an advanced stage of the cyber kill chain. This scenario is more likely investigated by a threat hunter rather that a SOC analyst.

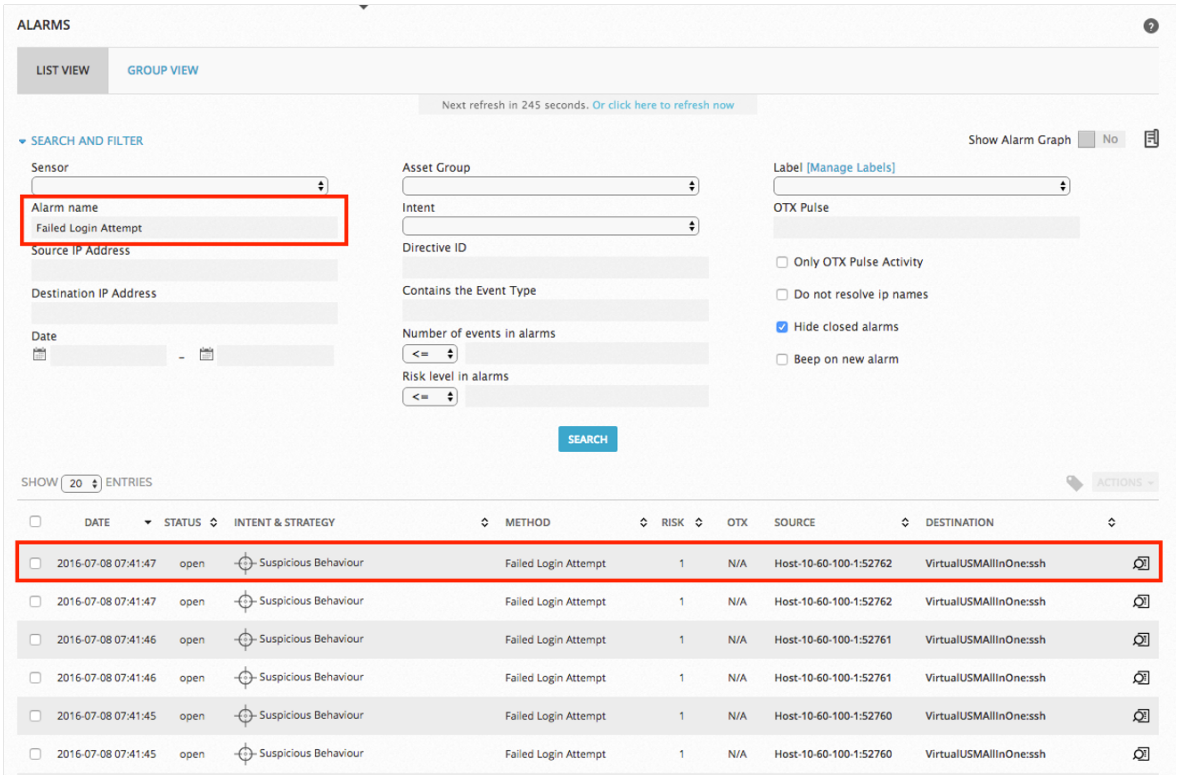


Figure A.6: scenario 4 - password spray

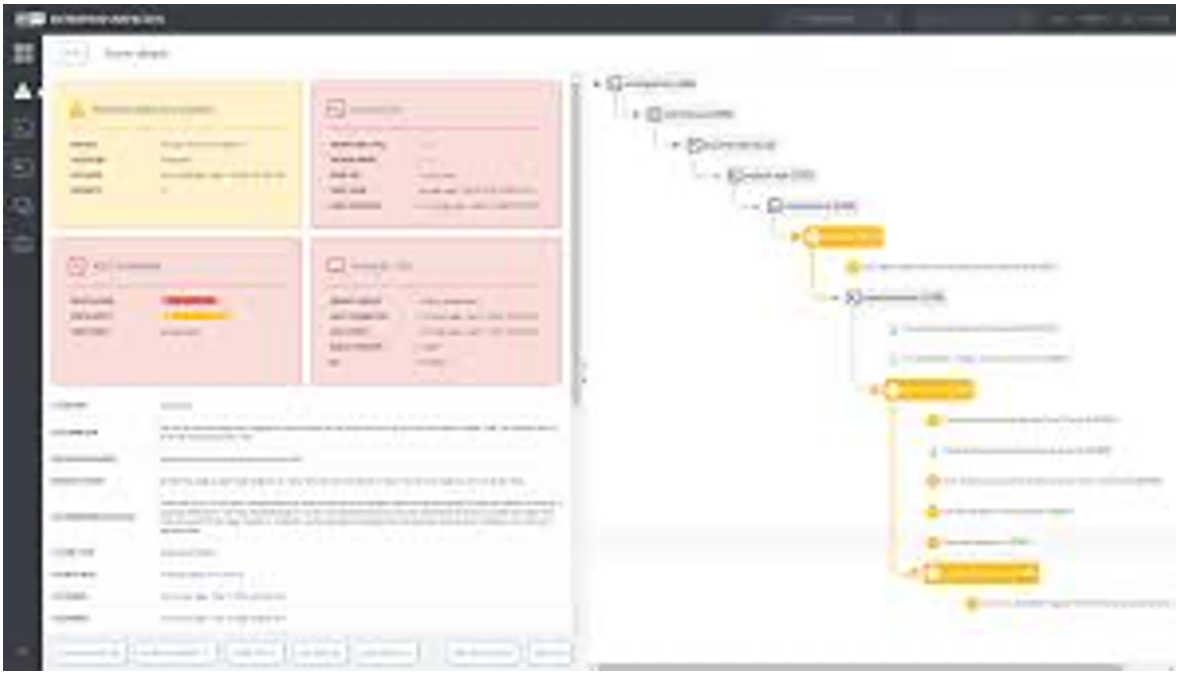


Figure A.7: scenario 5 - Advanced stage

Appendix B

Wargame

The following section contains the details of the wargame that was organized of for the second part of the research.

As was discussed in the methodology section the wargame was designed with many objectives in mind. The first one being observe the actions of the Red Team in a realistic scenario when the Blue Team is investigating their actions.

There are three elements that will be presented. The first one is the technology used to create the environment for the wargame. The second one is the design of the infrastructure and the planned attack paths. The third one is the actual outcome of the experiment and the recorded actions of the Red Team

B.1 Design

B.1.1 Infrastructure

The design of the infrastructure can be divided in two sections. The first one is the actual elements used to simulate a realistic organization network(servers, workstations, subnetworks, firewalls, etc.). The second are the tools used to support Blue Team activities. The details regarding this two elements will be discussed next.

Company network

The first step in the design of the company network was to plan the departments, their roles and relationships. The importance of this step is not be under estimated, because when the Red Team is exploring the network the decision they make is also based on the logical structure. There are four departments: IT, Sales, Engineering, Human Resources. The role of of each department is now explained. The members of the Engineering department use special software that require computational power and are therefore allowed to use remote desktop services for connecting to the company. The Sales department use specific CRM software to manage the customers, the software runs on an internal server which is accessible only to the Sales department. The IT is the biggest of the teams because it include many developers that work for the company. Finally, the HR department manage the employees and is able to reset passwords and add users to other departments. Figure X shows the relationship between the various entities.

SOC infrastructure

While the Red Team was free to interact with the system as they pleased, the Blue Team was monitoring and investigating their action the whole time. For this reason on each machine in the network was installed a sysmon. The specific sysmon configuration can be find at the link

SOC Console

Figure B.2 shows the basic SOC console of the SOC analysts. Figure shows an example of the workbook while being used during the wargame, it is possible to see how each event is already mapped to the ATT&CK framework. Figure B.4 is a view of a drill down performed during the wargame to gather more information on some suspicious event.

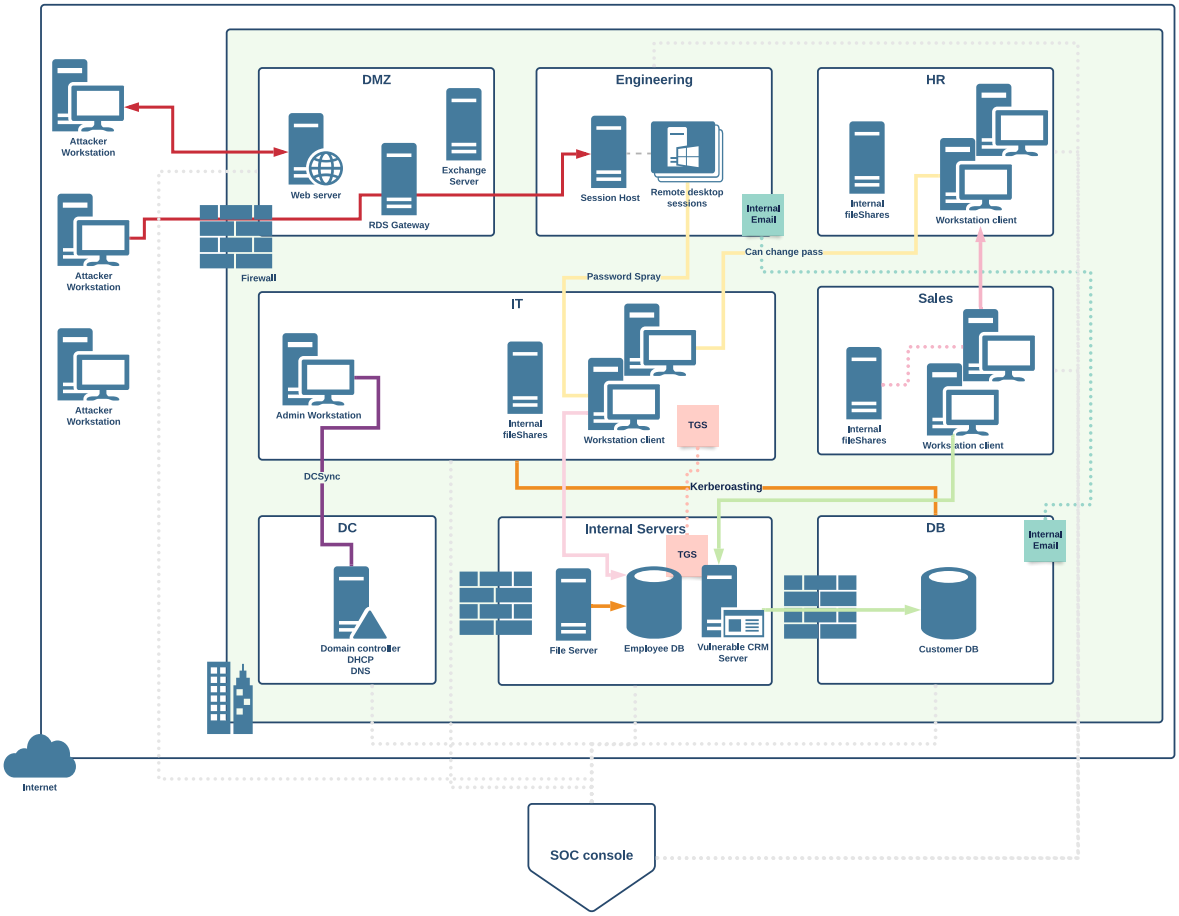


Figure B.1: Infrastructure design

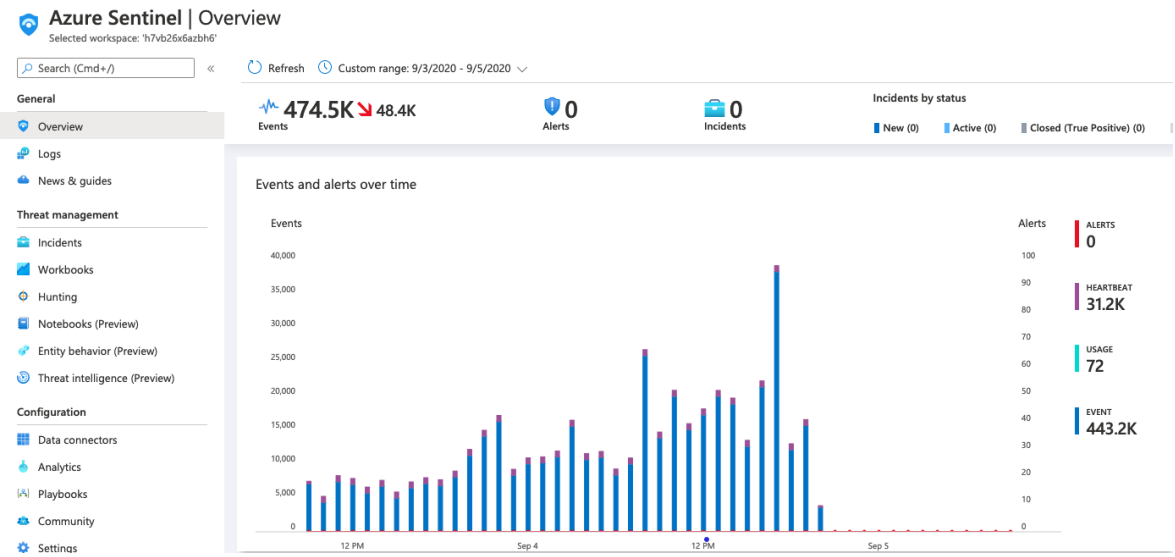


Figure B.2: View of Sentinel events

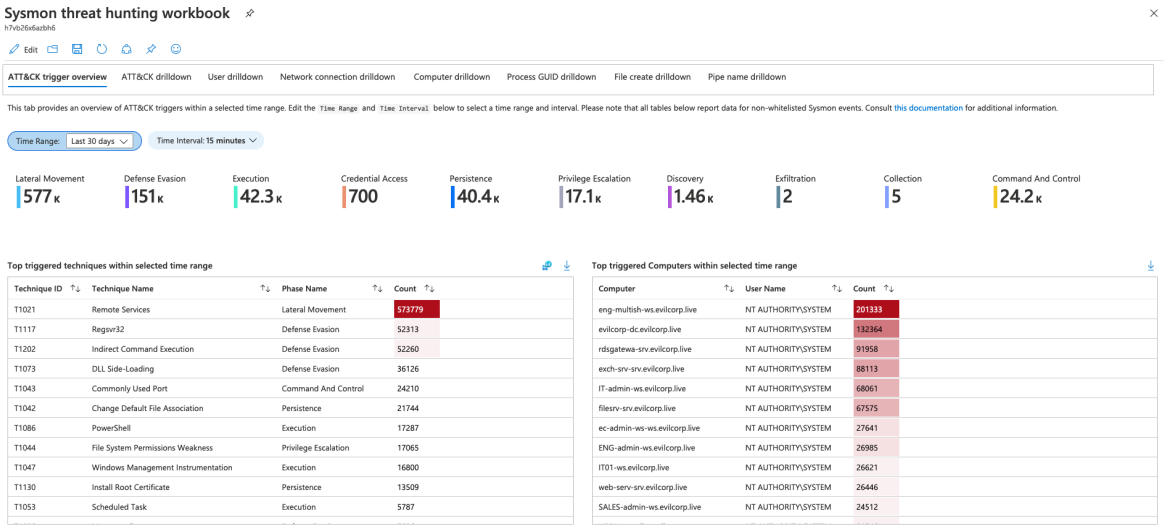


Figure B.3: Sentinel workbook mapping logs to ATT&CK techniques

B.1.2 Planned attack paths

Figure B.1 gives an overview of the scenario. The colored arrows corresponds to the various attack paths that were been planned.

Initial foothold path This path was intended for the Red Team to obtain the initial foothold. As can be seen in figure X the red path starts from the external webserver. The company is hiring new engineers and ask to the candidates to send their CV. The CV would have been received by the someone from the HR department which

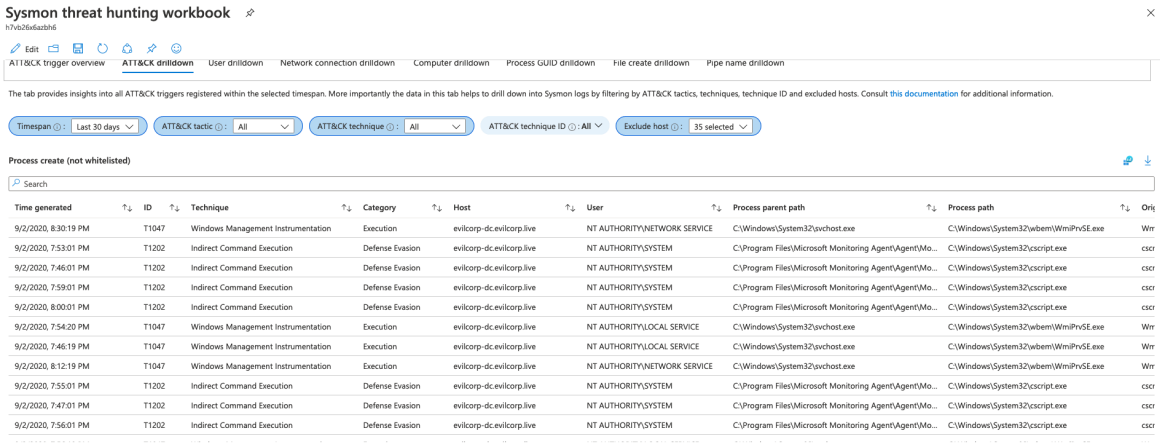


Figure B.4: Drill down in the Sentinel Workbook

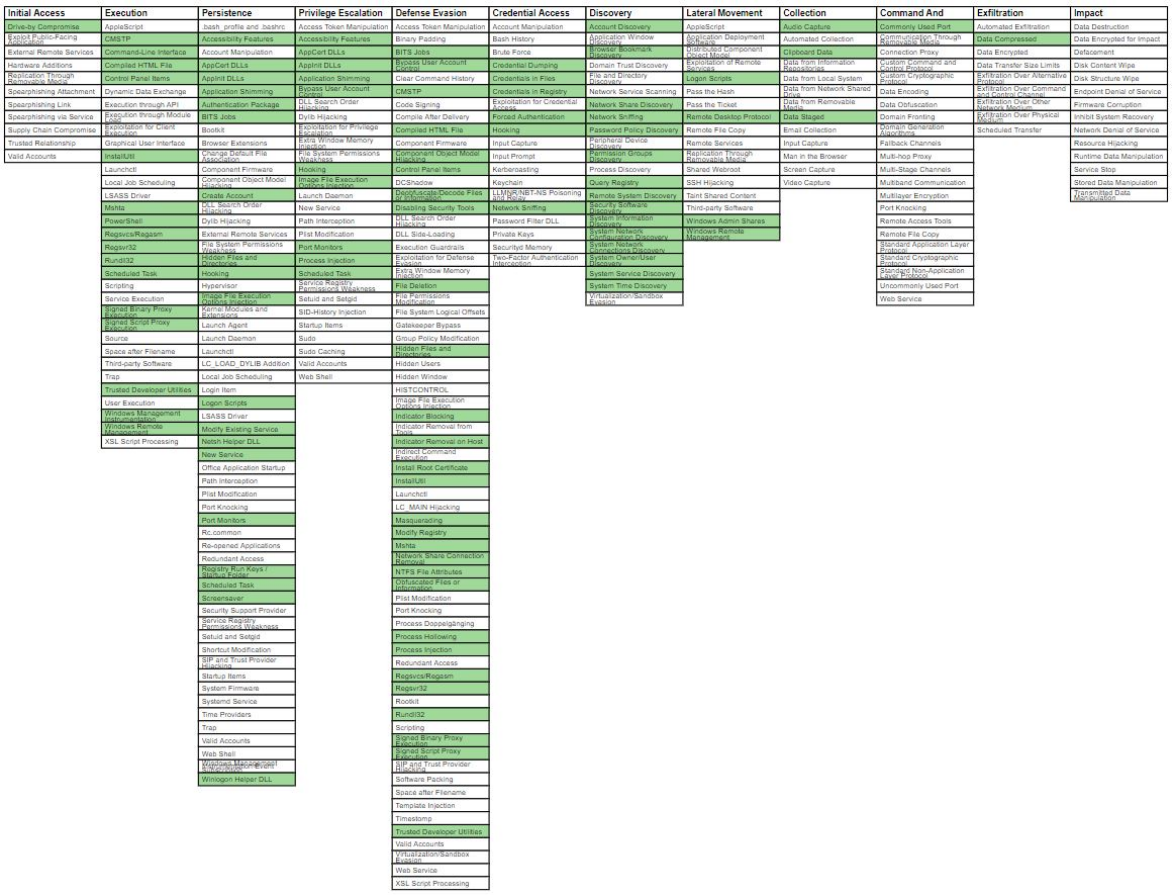


Figure B.5: Sysmon ATT&CK coverage - 156 techiniques

in turn would have forwarded it to the engineer manager.

In case the CV was the vector for a fileless attack the engineer manager would have run it from a remote desktop inside the company network, in the engineer department subnet. Alternatively, if the attack vector was a phishing website the

engineer manager would have provide the credentials for the Remote desktop services.

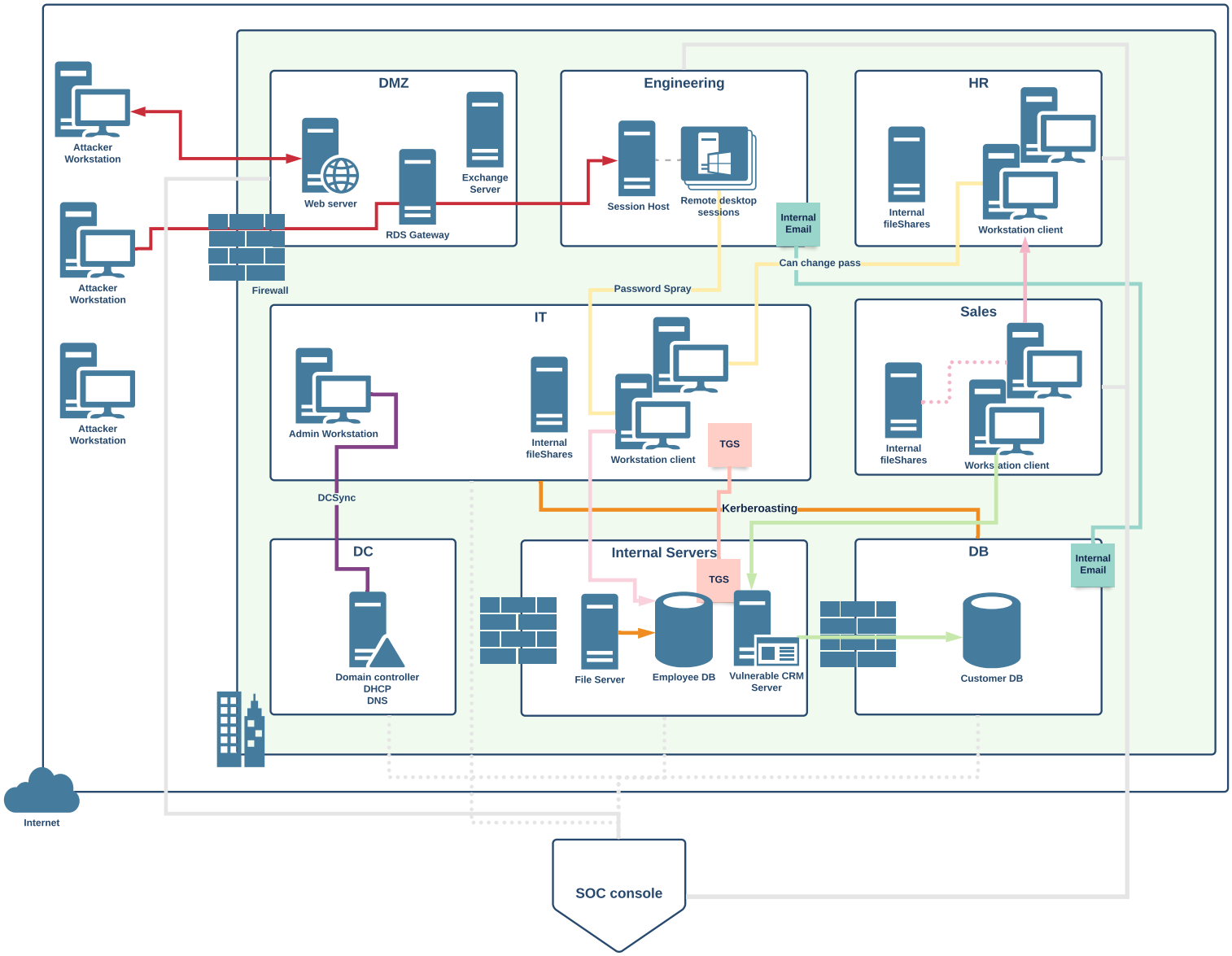
At this stage the attacker should have run Blodhound injestor to gather additional information on the network and should have figured out which attack paths could been followed.

Lateral movement 1 from Engineering From the computer of the the engineering user it is possible to perform a password spray attack. There are aproximately 200 users in the IT department, and about 5 of them have a very common password such as ”summer2020”.

Lateral movement 2 from Engineering From the computer of the engineering user it was possible obtain access to the sales administrator computer by using a derivative local admin attack. A derivative local admin attack consists of identifying users who have a

B.1.3 Unintended attack paths

For ease of maintance and fast trouble shooting a special user was added to the environment. The user was local admin to every machine and domain admin amongst many other high privileges. The purpose of this user, however, was to be used by the researcher and was not part of the planned scenario. The Red Team discovered this user and used it to obtain high priveleges without following the planned attack paths. A mistake of the researcher was not putting the user out of attack scope. This shortcut was used by almost all the participants to reach the crown jewels.


DMZ
Engineering
HR
Attacker Workstation
Web server
Exchange Server
RDS Gateway
Session Host
Remote desktop sessions
Internal fileShares
Workstation client
Internal Email
Can change pass
Attacker Workstation
Firewall
Password Spray
IT
Sales
Attacker Workstation
Admin Workstation
Internal fileShares
Workstation client
TGS
Internal fileShares
Workstation client
DCSync
Kerberoasting
DC
Internal Servers
DB
Internal Email
TGS
Domain controller
DHCP
DNS
File Server
Employee DB
Vulnerable CRM Server
Customer DB
Internet
SOC console