HALMSTAD UNIVERSITY

Digital Forensics, 180 credits

BACHELOR THESIS

Threat hunting, definition and framework

Theodor Liliengren, Paul Löwenadler

Examensarbete teknologie kandidat, 15 credits

Halmstad 2018-05-15

Threat hunting, definition and framework.

Theodor Liliengren & Paul Löwenadler

School of Information Technology, Halmstad University, Sweden, 2018

Acknowledgements

Many thanks to Magdalena Rosenberg and the IKEA OSS team for the opportunity. And special thanks to Jussi Jaakonaho.

Theo & Paul

Contents

List of Figures

List of Tables

List of Abbreviations

APD Advanced Persistent Defense
APT Advanced Persistent Threat
CMMI Capability Maturity Model Integration
F3EAD Find, Fix, Finish, Exploit, Analyze and Disseminate
HMM Hunting Maturity Model
IDS Intrusion Detection Systems
IOC Indicators Of Compromise
IPD-CMM Integrated Product Development Capability Maturity Model
IPDS Intrusion Prevention and Detection Systems
IPS Intrusion Prevention Systems
IR Incident Response
OODA Observe, Orient, Decide and Act
PDCA Plan, Do, Check, Act
RPD Recognition-primed Decision
SECM Systems Engineering Capability Model
SIEM Security Information and Event Management
SOC Security Operation Center
SW-CMM Capability Maturity Model for Software
SWOT Strengths, Weaknesses, Opportunities and Threats

Threat hunting, definition and framework.

May 15, 2018

Abstract

Being pioneers comes with advantages and responsibility. The concept of threat hunting is currently being subsidized by businesses promoting their products. Additionally, there is little or no information regarding the implementation and the effects, which vary depending on the organization. Threat hunting needed an unbiased definition in accordance with employees in IT security. Consequently, the frameworks used when assessing threat hunting had to be objective. This thesis presents a definition of threat hunting, composed using impartial opinions. Furthermore, the thesis provides unique frameworks to assist when implementing and assessing threat hunting at an organization. This thesis has several areas of application: as a knowledge base for threat hunting, as the recommended practice for implementing threat hunting and as groundwork for a more comprehensive evaluation of threat hunting capabilities. Ultimately, the thesis offers unprecedented nonpartisan information and recommendations on threat hunting.

1 Introduction

Today, individuals possess the means to be more powerful than ever before. Teenagers have single-handedly disabled air traffic control systems [60], shut down e-retailers [48] and manipulated trades on the stock exchange [32]. What motivated teenagers can do, well-funded organizations can also do, and probably better. In the IT world, people are viewed as threats. They are the adversaries. The greatest opponents to IT security are consistent, flexible and highly adaptable. The motivation and capability they possess enables them to seize every opportunity to cause damage [61] [17] [21]. They use sophisticated software to achieve their goals, and more often than not they use general IT tools rather than hacker-specific tools and malware [30]. The threats they pose are known as advanced persistent threats (APT's). APT's have the ability to initiate and maintain long-lasting operations against profitable targets. Highly motivated and well-funded attackers are not going to stop anytime soon, nor will threat hunters lay dormant until they strike. Hunters are actively trying to mitigate these threats and minimize or avoid the damage they can cause [35].

Attacks against businesses such as LinkedIn, Dropbox and Yahoo have resulted in millions of accounts and passwords being made available for anyone to download. Big corporations in America such as Target, Evernote and Living social have also been subjected to attacks, leaking over 100 million customer records in total [48] [24] [50]. This is causing companies to develop their own strategies for incident response. Incident response, also known as IR, is a collective name for detecting threats against the company and quickly implementing countermeasures. The corporation's level of defense is based on their resources and the current perceived threats.

Nowadays, incident response is very well developed and is used by almost all big corporations. For the most part, IR functions like a cyber fire department, meaning that the work begins only after there has been a breach. However, lately the focus has shifted from this reactive approach to that of a more proactive one. With the help of SIEM (security information and event management) systems [65] and automated analysis, security teams actively go hunting for flaws and adversaries in the system to prevent the breach from happening. The earlier you can detect and track your adversary's activities, the better chance you have of limiting or totally avoiding the damage they could cause [1] This type of proactive work is known as threat hunting.

A lot of professionals within the IT security business feel that threat hunting, at least partly, describes their current assignment. However, the purpose of threat hunting has never been about rebranding the work security personnel have been doing for years.

Instead, focus is placed on increasing proactive measures that are part of IT security operations.

Threat hunting requires a familiarity with the IT environment of the organization and also the analytical skills to form and examine your own hypotheses. The end goal is to automate as many analytical processes as possible, making the actual hunt faster, easier, more frequent and more accurate [35].

1.1 Problem definition

There is no such thing as an impenetrable network. An adversary will eventually figure out how to circumvent the protection that is in place, and inevitably, the network will become compromised. IT security has always been a game of leapfrog in which a threat is developed by an adversary and once it is discovered, security vendors will find a way to mitigate it. However, the biggest threats today are different. To steal information, disrupt services or conduct espionage on an organization, adversaries specifically tailor their cyberattacks to breach the target's protection. Therefore, IT security professionals are looking for new strategies to prevent these intrusions from happening and limiting the amount of harm inflicted by them.

Recently, the term threat hunting has emerged. Many claim that threat hunting is the way to tackle the advanced threats of today. Others claim that it is nothing new and that security personnel have been conducting threat hunting for many years already [18]. But what is threat hunting exactly? The majority of information available on threat hunting are written as short articles or blog posts, and the few scientific papers published are sponsored by the threat hunting industry. This makes the term confusing and difficult to grasp since everybody defines and implements threat hunting differently.

The purpose of this thesis is to investigate the current scientific field of threat hunting and the general process of threat hunting, to highlight opportunities and difficulties when combining these into a threat hunting strategy, but also finding suggestions on how this process can molded to avoid certain obstacles and make use of opportunities.

Thus, this thesis seeks to:

- Define threat hunting,
- Construct a framework for conducting threat hunting and assessing hunting abilities.

1.2 Goals

- To aid organizations by expanding knowledge and analyzing what the term threat hunting entails and how to implement it,
- To serve as groundwork for an evaluative study when considering investing in threat hunting.

1.3 Limitations

Although the concept of threat hunting is relatively recent, constructing new frameworks for conducting and assessing threat hunting would be counterproductive. Hence, this thesis will focus on investigating frameworks with features suitable for adaptation to threat hunting. The finalized frameworks will therefore include features from several different frameworks with additions and improvements tailored to threat hunting.

1.4 Problematization

Arguably the biggest hindrance with having multiple frameworks is that they are difficult to compose into one. Combining frameworks, trying to fit one framework inside another can prove to be problematic. To solve this problem one would have to be careful about what framework goes where. Modifications and compromises must be made to certain frameworks for them to be able to play along with the others. Different frameworks require different input in different situations. Being able to manage this, tailoring it so that the frameworks align and cooperate can be the difference between success and failure. As this paper aims to make use of multiple respectable frameworks to create a larger framework for conducting threat hunting, all the frameworks must function together.

Additionally, frameworks are hard to test, experiment with and explore. These difficulties will most likely increase as the main framework grows, possibly discouraging workforces from committing to trying a new approach to threat hunting.

Since threat hunting is a recent concept and the description varies depending on whom you ask, finalizing a definition that both affirm these conceptions and serve the needs of this thesis is necessary. However, the definition of threat hunting as presented in this thesis is not intended as exclusive or definite.

1.5 Ethical guideline

One purpose with this thesis is to aid organizations by providing a framework for implementing threat hunting. A problem with this is that an adversary could possibly use this to their advantage. If an adversary knows about the prevention techniques employed by their target it gives them a major advantage. Therefore, if an organization chooses to adopt the material this thesis provides they must be aware that it is public knowledge and accessible to anyone. However, since this thesis aims to produce prevention techniques that are independent of technology, there is very little opportunity for an adversary to gain an advantage by simply knowing about the frameworks this thesis presents

2 Related works

2.1 Threat hunting: Open season on the Adversary

The report *Threat Hunting: Open season on the Adversary* [18] written in April 2016 contains an analysis of the advantages of threat hunting as well as disadvantages of using current defense strategies. A poll with 494 participants showed that 86% of the people questioned were employed by corporations working with some kind of threat hunting. However, as many as 40% of those did not have a specific threat hunting strategy. Drawing conclusions from the result of the poll, one interpretation could be that companies still do not know how to implement efficient threat hunting. Currently, many organizations rely on known threats, manual analysis and tools which are not tailored to threat hunting [18].

Of the companies that implement threat hunting,

- 52% feel that threat hunting has helped them find previously unknown threats to the company,
- 74% feel that threat hunting has helped reduce the amount of cyber-attacks,
- 59% feel that threat hunting has increased the precision of their IR capabilities

Although the majority of the employees are positive towards threat hunting, most feel that there is room for improvement:

- 53% feel that their strategy is too obvious to adversaries
- 88% feel that their threat hunting tools need improving
- 56% are dissatisfied with the amount of time needed to perform threat hunting

The authors of the report conclude that many organizations experience enough advantages of threat hunting to invest in it. Unfortunately, many are still uninformed or unsure, requesting tools and strategies that are better adapted to employees [18].

2.2 Incident-centered information security: Managing a strategic balance between prevention and response.

Information security strategies are formulated using the basic principles of prevention and response. Prevention is the practice of managing anticipated threats while response is managing unforeseen threats. The importance of prevention methodologies are heavily emphasized in the strategies of current commercial organizations, but response strategies are equally important in the contemporary threat environment. The study conducts a case study to analyze how and why organizations balance between the two strategies.

As attacks are getting more and more sophisticated, organizations may need to reevaluate their balance between prevention and response strategies. Although focusing on preventive actions works well for mitigating repetitive and lowly attacks, the more sophisticated attacks require more focus on response capabilities. In addition to the rise in sophistication, there is growing motivation to attack information assets of commercial organizations. Such attacks have also become more lucrative, thus attracting organized crime syndicates. Furthermore, as information warfare tactics employed by nation states are evolving, anti-commercial attacks may be heavily financed and aimed at circumventing known preventive measures.

There is no perfect or correct strategy. Each organization makes their own risk assumption and evaluates their own threat environment. This affects the stance they take on balancing their prevention and response efforts. If an organization that has a highly unpredictable threat situation mainly focuses on preventive measures, the shortcomings in their response preparedness could lead to huge damage if they are subjected to an APT. In comparison, if an organization with a stable threat situation focuses too much on their response capabilities, they are likely to suffer frequent losses to well-known attacks. In both scenarios, the losses would not only be a direct result of the actual attack but also the inefficient use of security resources. However, there are situations when a balanced strategy is an inferior option, making security programs superfluous and more expensive than expected.

This study has developed an incident-centered framework to help evaluate an organization's current information security posture, thus providing awareness and knowledge about how they are balancing between prevention and response and why they prioritize the way they do [55].

2.3 Diamond Model of Intrusion Analysis.

Adversary, infrastructure, capability and victim; all malicious activity is related to at least one of the four mentioned above. Authors Sergio Caltagirone, Andrew Pendergast and Christopher Betz presents "Diamond Model of Intrusion Analysis" [8], a model that provides insight to threat analysis and -mitigation. The model was developed around the question asked by experienced analysts:" What is the underlying method to our work?" The four pillars of the diamond model are connected, illustrating their relationship as they together form the outlines of a diamond. The diamond model can easily be integrated in many frameworks, especially threat hunting, since it has a logical and easy-to-follow strategy for hypothesis generation. Sergio Caltagirone published several threat hunting strategies using this model. The diamond model also supports approach development, mitigation and hunting strategies, which is highly related to this thesis.

3 Method

3.1 Research

To define threat hunting, qualitative research will be conducted by first studying previous research about the phenomenon. A compilation of definitions made by the IT community will serve as the second part of the groundwork. Acquired material will be evaluated and the concluding definition formulated to be in accordance with the community. The definition will also serve as a foundation for the implementation of threat hunting.

To construct the frameworks, other relevant frameworks will be closely examined. Frameworks regarding implementation of new techniques, thought process, intelligence gathering, combat, quality control and general IT security will be highly relevant to analyze. Features considered adaptable to this thesis will be adopted and/or altered to fit its purpose. Features such as learning potential, flexibility and adaptability are of great importance. The finalized frameworks are going to address the conducting and implementation of threat hunting as well as assessing an organization's ability to benefit from threat hunting. Furthermore, the frameworks should allow for improvements.

3.2 Choice of methods

The choice of methods was greatly influenced by Jacobsen [29]. He states that qualitative research is most suitable when trying to clarify a phenomenon. Furthermore, Jacobsen suggests that an investigatory inquiry requires that the chosen method of research produces a nuanced result. According to Jacobsen, quantitative research is more adapt when investigating the scope of a phenomenon and does not lend itself well to nuanced observations.

The study will emphasize on the most reliable sources to make the result credible. Sources used in the research will be sorted by the following scale with descending reliability [66]:

- Scientific research, essays and journals,
- Course literature and similar material,
- Encyclopedias and lexicons,
- Government reports,
- White papers,

- The daily press,
- Popular science.

The results of research can be assessed by examining the internal validity [29]. The internal validity is asserted when the conclusion of the research is accepted as correct by the ones it is presented to. The internal validity of all conclusions this thesis presents will be subjected to this kind of verification. However, it is especially important to examine the internal validity of this thesis definition of threat hunting since there is no widely accepted definition of what threat hunting is.

3.3 Method critique

This thesis will mostly be based on information gathered from published literature. However, with threat hunting being a relatively new concept, the limited selection of research available presents a problem when only trying to include objective studies from reliable sources. The majority of information on threat hunting are written as short articles or blog posts, and the few scientific papers published are sponsored by the threat hunting industry. This problem can be tackled by widening the angle of information acquirement and critically examine the sources and legitimacy of documents. Sources will be considered using the method in 3.2. Lower tier documents can be used if they are supported by a higher tier source.

When conducting a literature study one will get many different opinions on the subject, thus providing a better understanding. Another benefit of a literature study is that it isn't limiting to a third party, unlike interviews, specific hours of the day or access to certain locations.

When literature and research theory is limited, the best suited approach is, according to Hsiu-Fang Hsieh and Sarah E. Shannon [25], a conventional content analysis (qualitative research). This method of research is conducted by absorbing as much data as possible to develop a complete understanding of a phenomenon, then analyzing the data repeatedly to gain new insights of it. This method is preferred as it uses previous research, meaning that the same experiments and work does not need to be repeated to achieve a similar result. A problem when using available scientific work is that it might not always be relevant and up to date, which may be less than optimal for a given situation. A secondary analysis can never as described in [12], be better than the literature that is used.

One downside to this approach is that the researcher might misinterpret the context of the data, thus the findings may not accurately represent the data [37]. This method of

research does not lend itself well to development of theories but the results can be used for concept development or model building [25] which in this thesis case is ideal.

Interviews can be a good addition to a literature study as it may provide knowledge on the field and information from companies, which makes the result more applicable. Interviewing comes with certain drawbacks too. Contact must be established and questions must be answered and processed from a subjective to objective view. The right people must be asked the right questions to get any valuable information on certain matters. Also, secrecy agreements between a company and an employee can limit the interview.

On a new subject like threat hunting, it is plausible that the interviewee, who might be an expert in his field, feels that the leading interviewer lacks knowledge to conduct the interview [4], which might affect the outcome negatively. For an interview to be used as a source it must be transcribed, so that the reader may take part of the conversation. Transcribing is time consuming, and majority of the work takes place after the interview [15].

Alternative method

An alternative method would be to map the origins of threat hunting and extensively explaining the evolution that spawned the phenomenon. This would involve researching the entire field of IT security and sequentially progressing through its timeline. While that method could produce a definition regarded as more accurate, it would be a huge undertaking and outside the scope of this thesis.

4 Theoretical Background

This section will describe a few topics that are highly relevant throughout the thesis and will in later sections be presented and discussed.

4.1 Intrusion Prevention and Detection Systems

The purpose of Intrusion Prevention and Detection Systems (IPDS) products is to monitor networks or systems for suspicions activity. Suspicious activity may include policy violations or activity with malicious intent. When potential threats are detected, the data is forwarded along with information related to the event. Related data may include the date and time the attack was detected, the type of attack and the source and destination IP addresses. Analysts use this information to manually inspect IDPS alerts. IPDS products use signatures to identify malicious activity. These signatures need to be updated on a regular basis to ensure that new attacks can be detected [10] [53].

4.2 Security Information and Event Management

Information generated from IPDS products are forwarded to a Security Information and Event Manager (SIEM) where alarms will be generated based on the analysis of the log data [10]. The SIEM is a hive mind for all relevant security data in the form of logs. It allows for an analyst to have all alerts viewed in the same user interface which will facilitate spotting trends and anomalies [52]. A SIEM system typically deploy a set of agents in a hierarchical structure. These agents collect security-related data from systems such as client devices, servers and network equipment. The data is forwarded to a centralized management system for inspection and this system then generates alerts when anomalies are detected. Figure 1 presents a typical example of a SIEM topology.

Modern SIEM systems must be able to upscale to handle the size of security-related traffic generated. Storage and the processing speed both have to meet the needs of the organization [22]. The systems must be flexible as well since data is generated from different types of devices throughout the network. Thus, organizations require for example distributed storage and a correlation engine that can process and enrich data sets at high speed. Nowadays, most of the leading technologies for handling large amounts of data are developed on MapReduce [54]. The framework MapReduce was designed by Google as a way of processing distributed data. When processing data with MapReduce, it performs two tasks simultaneously: While categorizing the input data into all possible statistics, it

is also running an analysis that merges the matching statistics into the answer with the highest probability. The implementation of this model allows for scalability by processing data over arrays of hardware. For further reference on MapReduce, Jeffrey Dean and Sanjay Ghemawat have written a paper called *"MapReduce: Simplified Data Processing on Large Clusters"* [13].

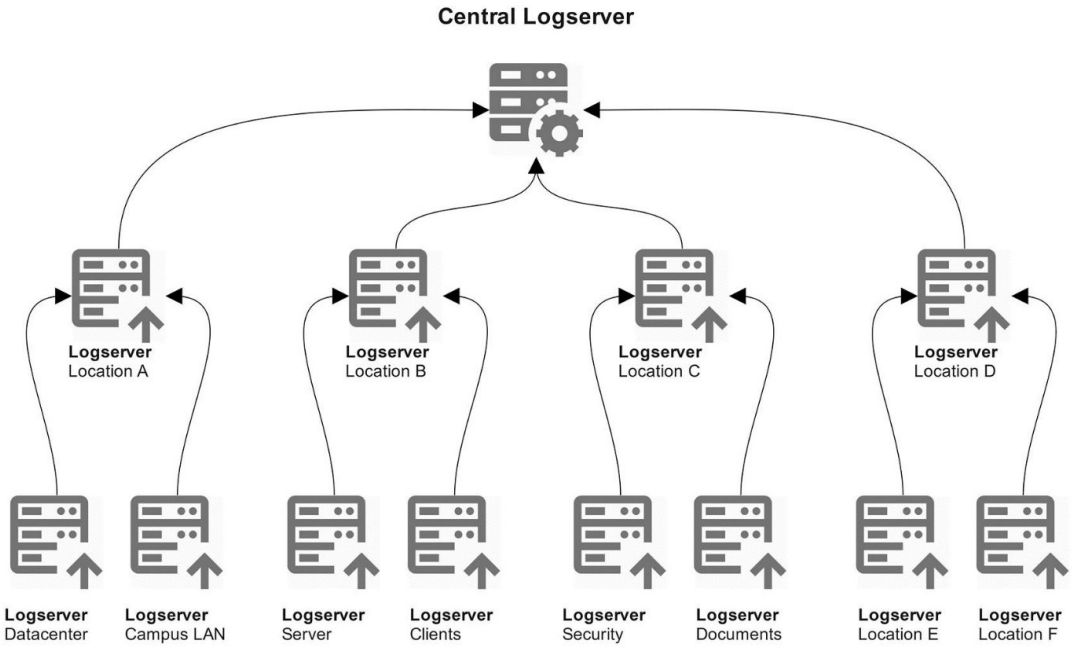

Figure 1: SIEM Topology

4.3 Antivirus and antispam software

Antivirus software is used to detect, prevent and remove malicious programs. If the software is kept up to date, the antivirus products currently available are effective at stopping many instances of basic malware. Anti-spam software is used to detect spam and prevent it from reaching users' email inboxes. Spam emails may contain malicious attachments or links leading to malware and other malicious content. Emails created with the purpose of deceiving its recipients and through them gain access to sensitive information are referred to as phishing emails. Alerts generated as a result of an increased amount of spam mails may indicate an ongoing attack [10].

4.4 Logs

Logs are recorded events from operating systems, services and applications and are of great value when an incident occurs. Important logs can be generated from VPN's, web servers, email servers, firewalls or custom internal applications, virtually everything that an attacker may target [52]. The logs, if configured correctly, will let you see which accounts were accessed and what actions were performed. Organizations should have a baseline level of logging on all systems and it should be more extensive on systems critical to the organization. Logs from network devices are often used when identifying network trends and anomalies. Endpoint logs are valuable when systems like laptops may exist outside network perimeter. Getting data directly from clients makes it possible for analysts to correlate internal activity [49].

4.5 Advanced persistent threats (APTs)

An APT may use techniques for intrusion such as spear phishing and social engineering to gain access to a target system. Unlike a simple cyber-attack where the attacker goes in and out swiftly to avoid being logged or countered, an APT is a long, ongoing operation where the attacker stays undetected by operating in a stealthy manner. This strategy is favorable when the purpose is to steal data over time, such as high value information. Once the attacker is inside, he will try to gather valid user credentials and start his lateral movement across the network, exploiting weaknesses to escalate his privilege and open backdoors. There are reports of APTs sitting in systems for months [43]. A good analyst should be able to detect an APT by analyzing anomalies in outbound data, as the attacker surely will try to send data back to base [51].

4.6 The cyber kill chain

Originating from the military, the kill chain (represented in figure 2) was used as a concept to describe chain of events during the enemy's operations. Recognizing these stages in an attack is the key to breaking the chain, thus neutralizing the immediate threat. The cyber kill chain is a framework created by Lockheed Martin and describes the steps taking place in most cyber intrusions or APT's. It revolves around intrusion, since this was the focus of cyber security when the framework was created [19].

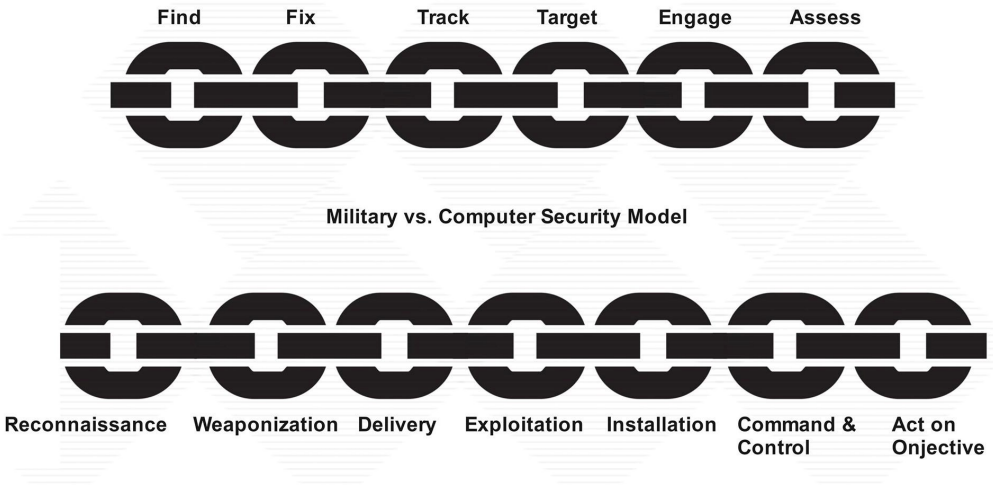

Figure 2: The cyber kill chain

4.7 Threat Intelligence

Cyber Threat Intelligence is information, or often referred to as intelligence that generally will help organizations to fathom the current cyber threat situation [26]. Before threat intelligence has been analyzed, interpreted or enhanced it is referred to as threat information, which is a more raw and unprocessed state. Cyber threat information is any information that can in any way ease an organization ability to identify, assess, monitor, and respond to cyber threats [31]

Some examples of threat information include [31] [46].

- Indicators (system artifacts or observables associated with an attack).
- Tactics, techniques and procedures (TTP's).
- Security alerts.
- Threat intelligence reports.

- Recommended security tool configurations.

Threat intelligence often come in the shape of reports and are constantly created as the threats evolve and change. These reports are shared within an organization or even better, on the Internet [39]. These procedures and common to be part of the general IT security practices. This allows network defenders worldwide to take part of the current threat situation and create a better situational awareness [31] [39].

5 Defining threat hunting

Threat hunting is a big but confusing topic in cyber security. Therefore, we need to define threat hunting, establishing a foundation for this thesis. The definition will be formulated to correspond with the perception within the cyber security community. Here are some examples of how they define threat hunting:

- "Threat hunting is the process of seeking out adversaries before they can successfully execute an attack." [38].
- "Threat hunting can briefly be described as proactively seeking indicators of compromise" [36].
- "While tools are important, threat hunting is not specific to any technology nor is it dependent on them. Rather it is about knowing when, where, and what signs to look for" [63].
- "It's a more open-ended action where hunches, gut-feelings, and general security and risk-based experience drive individuals to places and activity they should analyze" [63].
- Threat Hunting is a strategy that begins with the assumption that the organization has been breached, and working backward from there to either detect the source, or to make sure there isn't an attack [63].
- "Cyber Threat Hunting refers to proactively and iteratively searching through networks and datasets to detect threats that evade existing automated tools" [58].

A survey made about threat hunting is highly relevant information. Reading the report of the survey helped interpret what the author and the ones questioned feel is important in threat hunting. Additionally, a brief comparison to related strategies will map the outlines of the definition, where threat hunting starts to bleed into other areas.

5.1 Summary of a threat hunting survey

Threat Hunting: Open season on the Adversary [18] written in April 2016, is a report based on a survey performed by SANS to examine the current state of threat hunting within organizations. The survey had responses from 494 participants. Of these, 22% where companies with 1001 to 5000 employees, 20% where companies with more than 50,000 employees, 18% where companies with 100 to 1000 employees, 17% where companies with 10,001 to 50,000 employees and 12% where companies with 5001 to 10,000 employees. Companies in the financial service business was the largest group of those questioned (22%), with government, high tech and education coming in as the second, third and fourth largest. Of the respondents, 29% were security analysts, 15% were security managers or directors and 13% were incident responders [18].

5.1.1 Current state of threat hunting

The survey reveals that organizations are unsure of how to develop a threat hunting program, especially forming the threat hunting team and automating their processes. This is not unexpected since there is no recognized definition of threat hunting. Nevertheless, a large part of the respondents still claim they practice threat hunting (see figure 3). *"For organizations that are performing threat hunting, less than 3% follow any formal, published, external methodology [18]."*

Figure 3: Does your organization perform threat hunting (Recreated from Threat Hunting: Open season on the Adversary)

When asked if their organization perform threat hunting, the majority (52.5%) answered yes. However, their threat hunting operations are conducted on an ad hoc basis. These organizations mostly let already familiar indicators of compromise (IOCs) initiate the threat hunting and almost exclusively utilize manual analysis to accomplish it. According to SANS, this reactive approach to threat hunting is inefficient. They state that to be efficient, organizations should strive to be able to launch hunts automatically and frequently without needing to first observe an IOC.

In terms of formal methodologies, threat hunting is still very much under development. To monitor the success of their threat hunting operation, SANS states that organizations should observe these three factors:

1. **Dwell time** - how long is the adversary in your organization?
2. **Lateral movement** - how much damage is the adversary causing, in terms of how many systems are compromised?
3. **Reinfection** - how many times has your organization been compromised by the same adversary or the same threat?

If no improvements are shown in these three areas, the organization needs to reevaluate their hunting operation. Moreover, SANS claims that the success of a threat hunting operation is having a process based on a well-defined methodology with continuous and frequent hunts. The more frequently an organization performs threat hunting, the less damage can be accomplished by an adversary. The survey reveals that 38.2% of the organizations consider themselves to hunt continuously and 34.3% hunt on demand. The rest of the respondents say that they either hunt on a regular schedule (15.5%) or infrequently (12.5%). Furthermore, 56% are unsatisfied with how long it takes them to hunt for threats which suggests that they are still trying to optimize their operation.

When asked if threat hunting provides value to the organization by a measurable risk reduction 52% answered in the affirmative. Apparently, the benefits of having a threat hunting program are acknowledged by many, especially by those involved with IT security. However, the survey shows that their organizations are unwilling to invest in a formal threat hunting program. Less than 30% of the respondents have a designated program with assigned staff, around 23% run hunting programs which draw staff from other IT operations while the largest group, around 40%, does not have a formal threat hunting program at all. Though, organizations are aware of their shortcomings as 88% of them feel that their threat hunting abilities needs improving upon.

"Overall, threat hunting is recognized as adding value, but organizations need to continue to develop their threat-hunting methodologies [18]."

5.1.2 Staffing and skills

When investing in security programs, organizations normally prioritize spending money on technology. Although organizations still prioritize technology in the case of threat hunting, they seem to recognize the value of investing in competent staff as well. For people who have experience from security operations centers (SOCs), starting to work with threat hunting is a natural transition. However, their qualifications are highly sought after, preventing companies from transferring them from their current positions. Therefore, investments in training other IT personnel in these new skills have increased.

SANS also requested that organizations listed what skills they valued in their threat hunting personnel. Incident detection, incident response and forensic analysis ranked highest amongst the skills mentioned. Unsurprisingly, the common denominators of these skills are uncovering evidence of either a threat or an intrusion.

5.1.3 Methods of threat hunting

SANS suggests that there are two different methods to threat hunting: network- or host-based hunting. When performing network-based hunting, you search for IOCs by logging and examining the network traffic. Host-based hunting requires extensive information about the system being analyzed. You need to know what is installed on the system and how the system is supposed to behave. Then you analyze how it behaves, what is installed and what processes are running to find signs of compromise.

When performing network-based data collection, one sensor can collect data from thousands of systems at once, making it much more scalable than host-based collection. Unsurprisingly, the top three collections utilized when hunting (network artifacts and patterns, IP addresses and DNS activity) are all collected from network traffic. Host-based data is typically collected via agents on each system and is generally more time-consuming. Also, different hosts require different analyzing strategies because the typical conduct of each host differ.

5.1.4 Tools used for hunting

Currently, new tools tailored to threat hunting are available on the market. However, 87% of organizations are using preexisting tools to discover IOCs. Furthermore, 67% are using vendor-provided or open source tools to perform hunts. Based on the immaturity of the threat hunting market, these results are as expected. The use of open source tools is often linked with ad hoc threat hunting capabilities and hunting operations lacking adequate funding.

According to SANS, existing tools provide valuable information about the flaws in your current capabilities. However, once the flaws have been recognized, organizations should develop customized software to increase the effectiveness of existing tools. When this cannot be accomplished, organizations should invest in third-party hunting tools to enhance their hunting abilities.

5.1.5 Covering your tracks

Threat hunting is ineffective if an adversary knows how you are tracking them. One of the characteristics of a mature threat hunting program is the ability to deploy undetectable operations. Unfortunately, only 23% of respondents consider their hunting processes to be invisible to adversaries.

5.1.6 Automation

SANS views automation as essential to successful threat hunting. Adversaries regularly employ automation to launch attacks against large numbers of organizations. This is one explanation why they are so successful. Automation is a great tool for exceeding the capability of the human brain. SANS believes that in order to combat threats, threat hunting programs need to utilize automation. They claim that although threat hunting cannot be 100% automated, automation should be used for time-consuming and repetitive tasks to uncover IOCs. Consequently, employees can focus on investigating and analyzing the high-priority IOCs to determine the magnitude of the threat that they pose.

"To put it another way, computers would analyze large amounts of information that have a low probability of an attack, while humans analyze small amounts of information that have a high probability of an attack [18]." Results of the survey show that 36% of respondents have reached an automation level of 51% or more [18].

5.1.7 Data needed for hunting

86 % of respondents regard anomalies to be the most common trigger for launching a hunt. An anomaly is anything that differ from the normal manner of conduct. But the results also show that organizations monitor external sources to determine what to hunt for. External intelligence provided by peers, media and intelligence providers was chosen by over 60% to have a major influence on their threat hunting procedures.

As previously mentioned, anomalies in networks and endpoints have the greatest impact on the threat hunting conduct of an organization [18]. A prerequisite to detect anomalies in your system is a set baseline of normal behavior. Therefore, the collection of data from all reporting systems is crucial to the threat hunting program being successful. The most valuable data feeds, as recognized by more than 60% of the respondents, are IDS/IPS feeds, access and/or authentication logs, DNS, network traffic flow, endpoint security feeds, SIEM alerts and threat intelligence sources. The quality of the data leading the program vastly affects the outcome of the hunts. The data should be validated regularly. If a certain type of data proves itself effective, more of it should be collected. Subsequently, if data being collected is not producing any result it should be removed from the collection process. The collection should be modified based on the value the data provides.

5.1.8 Analyzing

The majority of those questioned are using basic searching techniques to discover IOCs. These techniques will become less useful as the basic points of compromises are identified. Mature threat hunting operations utilize more sophisticated methods of analysis such as statistical analysis, sample aggregation and machine learning. The more complex the methods of analysis are, the less organizations choose to use those methods.

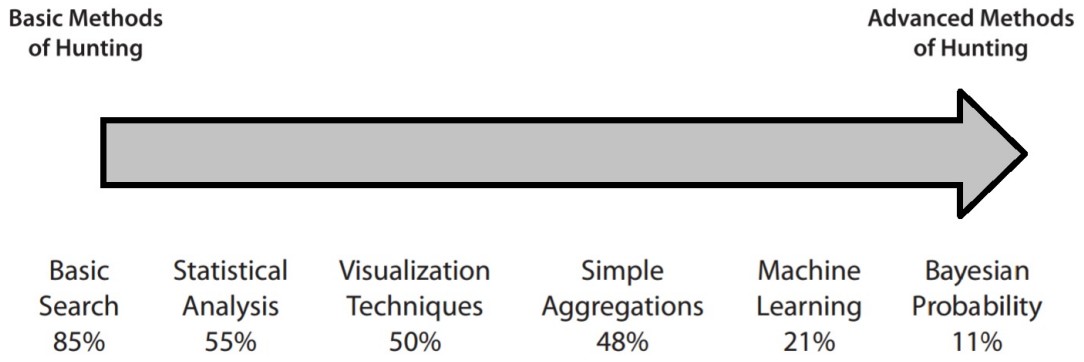

Figure 4: Methods of hunting (Recreated from Threat Hunting: Open season on the Adversary)

"This figure (figure 4) provides a guide to threat-hunting maturity and the next steps organizations should take as they progress toward more mature and advanced hunting capabilities"[18]

5.1.9 Results

As previously stated, 52% of participants feel that they are seeing a measurable risk reduction as a result of practicing threat hunting. The risk reduction is being observed as reduced attack surface, reduced exposure, improvements in security systems, increased speed of containment and fewer actual breaches.

Another indicator of success of a hunting program is the speed with which adversaries are being detected and hindered. Although organizations are seeing more adversaries caught with threat hunting, the harm being inflicted is still unacceptable. A large group of the respondents say that it takes them between one and eight hours to detect and respond to an adversary. However, according to SANS these claims are extremely optimistic as in reality, because of the immaturity of their threat hunting programs, organizations have no clue how long they have been compromised [18].

5.1.10 Improvements needed

88 % of organizations are seeking improvement to their threat hunting capabilities. For instance, 58% feel that they need better detection, [45]% needs more automated tools and 54% request better qualified personnel.

Many of the respondents launch their hunts in response to an anomaly [18]. They then present the anomaly to competent analysts to get an understanding of what happened. This approach is not very efficient. Instead, organizations should invest in tools and automation to improve detection and allow personnel to analyze more events, thus catching more adversaries [18]. The survey data shows that threat hunting is something organizations are willing to invest in. 62% are increasing their spending on threat hunting this year and over 42% are willing to increase it by 25% or more.

5.2 Conclusions on the SANS survey

Organizations feel that conventional security measures are failing to detect adversaries, both those trying to get in and those already inside their network. Therefore, they are looking into threat hunting. The vast majority of participants feel that their organizations are currently performing some type of threat hunting. However, they are still trying to figure out what a threat hunting program should look like. Presently, most of the threat hunting is done on an ad hoc basis and organizations do not have the level of automation needed for a successful threat hunting operation [18]. Nevertheless, the value of threat hunting is acknowledged by the organizations, both for finding new threats and discovering flaws in their existing security systems. Participants seem to agree that threat hunting has helped reduce the overall risk to the organization [18].

The respondents believe that it is important to evolve the threat hunting program at their company to present improvements, thus maintain the support of the leadership. The majority of organizations partaking in the survey express a willingness to invest in threat hunting.

Among the most valued skills for threat hunting are experience in incident detection, incident response, forensic analysis and penetration testing. This is not surprising since those tasks require you to be able to find pieces of evidence that are hidden on a system or cannot easily be found.

5.3 Comparison

The following comparisons seek to highlight both differences and similarities between threat hunting and conventional practices.

5.3.1 Incident response

If an organization's incident handling program is to be successful, preventive actions need to be taken to keep the number of incidents as low as possible. A reactive approach to incident handling means only responding after being notified of a breach. The incident response begins as a reaction to the incident. Threat hunting advocates a proactive approach to incident handling. Ideally, appointed threat hunters would actively go looking for threats based on intelligence, anomalies or even suspicions. Once a threat has been found, the severity is assessed and the appropriate response is conducted. A proactive approach has its obvious benefits, but allocating the incident response team to threat hunting can be very expensive. Also, as previously mentioned, surveys suggest that few organizations have an established threat hunting procedure. However, the proactive and reactive approaches are not mutually exclusive and an organization will need to have an IR team ready to handle the threats unveiled by the proactive work. Some proactive strategies can be found in documentation about incident response; even though they are two different approaches, they share some common ground.

These are some examples from the computer security incident handling guide, published by NIST [10]:

Systematic risk assessments should be a part of the regular routines of the organization. These assessments identify and estimate risks within the organization's systems and applications. After having identified a risk, you can either mitigate, transfer or accept it. Additionally, risk assessments help locate critical resources, which can then be prioritized and receive increased monitoring.

Hosts in the network should be configured properly. Applying the principle of least privilege ensures that each host is only granted the authority necessary to perform their assignments. All software used should be kept up to date and unused software should be removed. Furthermore, hosts should be constantly monitored and log all noteworthy security-related events, and all parts of the IT infrastructure should employ some kind of anti-malware software.

Configuring the network perimeter to secure all external connection points such as virtual

private networks significantly decreases the risks of security breaches.

Finally, organizations should educate employees about the correct use of its IT equipment. Examples of past incidents should be acknowledged, thus reducing the risk of reoccurrence. The organization must also ascertain that the IT personnel receives the proper training to maintain a high standard of IT security.

5.3.2 Incident detection

Incident detection refers to all the processes involved in detecting a threat, whether it is before, during or after an incident has occurred. Threat hunting is the spearhead in incident detection, focused on detecting and mitigating threats to reduce the risk of exploitation. Since incident detection is a large part of threat hunting, they utilize similar tools and strategies.

5.3.3 Forensic investigation

One part of threat hunting is to determine whether and how certain systems can be exploited. This differs from forensic investigation, where the systems examined already have been exploited. While forensic investigations try to obtain evidence of malicious actions, threat hunting also aims to find evidence of potential future breaches. Although similar mindsets are applied when analyzing systems, forensic investigation, unlike threat hunting, have plenty of reliable, repeatable and well-documented methods for procedures.

5.3.4 Penetration testing

Penetration testing has ties to both incident response and threat hunting. What is currently partly referred to as threat hunting was previously known as ghost level red team or ghost level penetration testing, a simulation exercise where the defenders try to find the intrusions [34]. Red teaming is an advanced type of assessment, used to identify weaknesses in security systems. Seeking venues to successfully penetrate the opposing team's defenses is proactive and can be compared to threat hunting mindsets, where you try to think like an attacker [62].

5.3.5 Threat intelligence

Threat hunting can be viewed as an extension to threat intelligence, A threat hunter might see threat intelligence as a start for hunting, and not an end result. As of today, we find that is no substitute for human analysis, meaning threat intelligence is perhaps not enough, but indeed a good foundation to threat hunt of off [39]. Threat intelligence can be combined with automated solutions and human data analysis to create a more organic form of threat detection.

5.4 Conclusions on the comparison

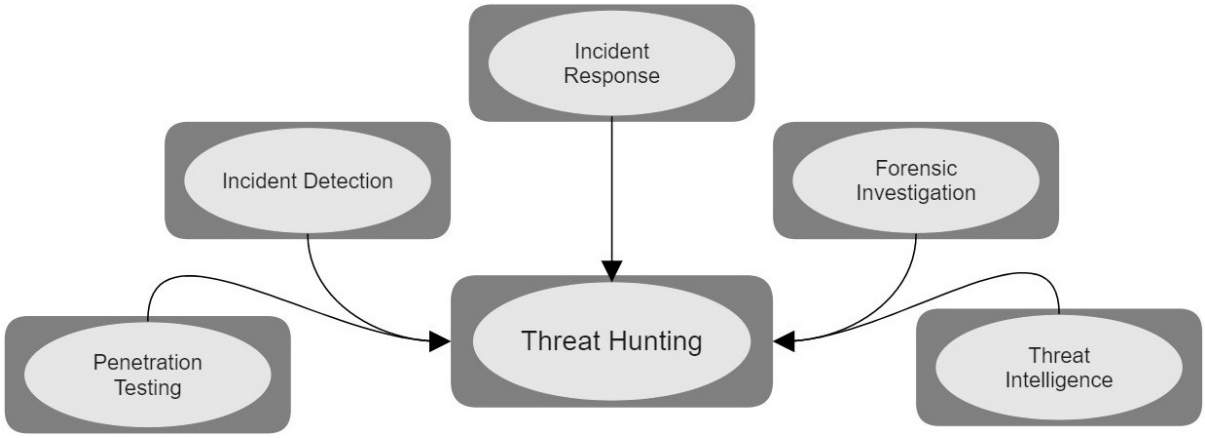

Figure 5: Related concepts

Concepts like threat hunting do not magically appear; they have some sort of origin. It is obvious that many techniques used by a threat hunter are already familiar. Incident response and detection, forensic investigation, penetration testing and threat intelligence are all considered related to threat hunting in this paper as illustrated in figure 5. There are other practices that could be relevant, but these were chosen from the results of the SANS survey [18].

5.5 Definition used for thesis

Threat hunting is neither a tool, nor a product. Threat hunting is a creative expression of data analysis to identify weaknesses, threats and ongoing attacks by proactively searching for indicators of compromise that has not yet been flagged by various detection technologies.

Whether a new way of thinking using familiar techniques and tools or a buzzword backed up by the big vendors in the threat hunting business, it is an approach that has been proven to benefit many organizations [18]. Threat hunting has a reputation of being abstract and artful. It could be defined as a creative expression of data analysis to find things bypassed by other tools.

6 Minor pre-study for assessing an organization's hunting abilities

In this chapter, a method for assessing hunting abilities is presented. To assist in concluding if threat hunting will benefit the organization, it is important to have an understanding of what qualities and resources an organization possess. This assessment is independent of software and will focus more on the characteristics of the IT infrastructure.

6.1 SIEM solutions

Relevant data is needed to hunt, but there are challenges with large-scale data collection and processing. Organizations need to allocate adequate resources both to collect data and to serve it to analyzers in an efficient way.

There is no single log standard. Therefore, not all products are compatible with the organization's logging requirements. This is important to consider when choosing the right solution for the organization. Furthermore, the SIEM should provide a way to integrate new logs since the organization might need to add logs from unknown sources in the future.

Currently, there are multiple SIEM solutions on the market. When choosing a solution for the organization there are various features and characteristics to consider. Vendors license products differently. Organizations should know the storage capacity needed to store their logs and choose a license that accommodates for that amount of data. Future growth is also important to consider, especially in the storage department [2].

The correlation of data is critical for threat hunting, and the correlation engine of the SIEM system should be able to cross reference from multiple systems and timeframes. Another highly important feature is automated response. The operators of the SIEM should be able to configure it so that when a threat is detected by log correlation, the system can contain or prevent it with suitable response actions automatically. If an organization chooses a product which can correlate data with ease and have the ability to respond to threats automatically, it will most likely see positive return of investment because the man hours required to analyze data will decrease.

Since the SIEM will be the main tool when performing a threat hunt, it is important that the user interface of the product is well designed so that it does not hinder the productivity of the operator. Ideally, tools for customized reports should be built into

the software to aid in spreading the findings and result of the hunt. SIEM systems require lots of resources to work efficiently, but if equipped with sufficient hardware and software and with proficient staff, most SIEM solutions will scale to the needs of any organization.

6.2 Capability Maturity Model Integration (CMMI)

According to the paper "CAPABILITY MATURITY MODEL INTEGRATION" [11], many models for improving the processes of organizations and companies have been developed. However, these models are too specific and lack a systematic approach. The CMMI is a generic process model that can be used to measure an organization's maturity in product and service development. It also serves as a framework for organizations and companies to make the proper investment in order to reach their desired maturity level. CMMI is a method developed at Carnegie Mellon University in 2005 and managed by the Software Engineering Institute. It is a combination of three separate models:

- Capability Maturity Model for Software (SW-CMM) [44]
- Systems Engineering Capability Model (SECM) [3]
- Integrated Product Development Capability Maturity Model (IPD-CMM) [40]

CMMI can be used to evaluate, certify and improve the quality of processes in organizations. Most importantly, the CMMI provides maturity or capability levels that are identified by a number. The maturity levels are designed sequentially like steps on a ladder, and a level cannot be skipped.

"Moreover, the model is conceived as a core, onto which further extensions can be added." [11]

Table 1: Capability Maturity Model Integration

Identifier	Capabilitylevel	Description
0	Incomplete	Incomplete
1	Performed	Processes are unpredictable, reactive and poorly controlled
2	Managed	Most processes are specific for certain projects and are mostly reactive
3	Defined	Processes characterized for the organization and are mostly proactive
4	Quantitatively Managed	Processes are measured and controlled
5	Optimizing	Focus on improving existing processes

6.3 Hunting Maturity Model (HMM)

The threat hunting company Sqrrl has used the CMMI as a core for their model used to determine an organization's hunting capability and maturity levels. This is a part of both Sqrrl's and SANS's framework for threat hunting. In the white papers *"The Who, What, Where, When, Why and How of Effective Threat Hunting"* [35] and *"A Framework for Cyber Threat Hunting"* [57], we can see this adaptation of the CMMI explained. The HMM is developed by David J. Bianco, and consists of five different stages of maturity or capability to measure and pinpoint to what extent an organization has implemented threat hunting.

Table 2: Hunting Maturity Model

Identifier	Identifier	Description
0	Initial	Relies primarily on automated alerts. Has little or no routine data collection.
1	Minimal	Makes use of Threat Intelligence for searches. Has a moderate or high level of routine data collection
2	Procedural	Follows data analysis procedures created by others. Has a high or very high level of routine data collection
3	Innovative	Able to generate new data analysis procedures. Has a high or very high level of routine data collection
4	Leading	Most data analysis procedures at this point are automated. Has a high or very high level of routine data collection

6.4 SWOT analysis

When deciding if threat hunting is a suitable strategy, a SWOT analysis can prove useful. Solely implementing threat hunting without a thorough analysis can cost any company unnecessary resources. A SWOT analysis will show who can benefit from threat hunting and why an implementation might be favorable. The SWOT analysis is an effective way of identifying strengths, weaknesses, opportunities and threats an organization might encounter [42]. Figure 6 represents the preformed SWOT analysis.

Figure 6: Our SWOT analysis

6.4.1 Strengths

Incident response, centralized logging, network and host visibility are four functions that should already be well integrated in an organization when researching the viability of threat hunting. The ability to react to threats is crucial to any organization that wants to implement threat hunting or for the incident response team to evolve into threat hunters. Centralized logging and visibility is part of how the incident response team will know when and how to act.

6.4.2 Weaknesses

Threat hunting can be used to compensate for weaknesses in an existing IT environment by detecting threats and weaknesses early. Actively going through the process of analyzing data and looking for compromise will reward the security operation center (SOC) with new experiences and knowledge. Because tools already used by the organization for cen-

tralized logging also can be used in threat hunting, implementing threat hunting does not initially mean a rework in the budget. However, a threat previously undetectable might need additional mitigation tools which then have to be acquired by the organization.

6.4.3 Opportunities

Hunting will eventually lower the overall risk of compromise at the organization as well as improve detection and response capabilities within the security team. Reports and debriefs can be used to build new detection systems and strategies. In addition to threat intelligence, threat hunting can increase the business intelligence, meaning hunters will learn more about their own network.

6.4.4 Threats

In the paper "Organizations and Cyber-crime: An Analysis of the Nature of Groups engaged in Cyber Crime" [7], the authors state that it requires an exceptionally closed-minded person to deny that sovereign states are capable of criminal acts. The disclosures of Edward Snowden confirm the accusations of Russia being part of distributed denial of service attacks, the Chinese authorities being engaged in industrial espionage and that the United States' government are engaged in worldwide cyber-surveillance.

As stated in the introduction of this thesis, individuals today have the potential to be more powerful than ever before. Teenagers acting alone have disrupted air traffic control systems [**1**], shut down major e-retailers [48], and manipulated trades on the NASDAQ stock exchange [32]. Although these cases are unique and particularly difficult to execute, it does not take a genius to perform a cyber-attack.

Insider threats come from within an organization and are either maliciously or non-maliciously intended.

6.5 A measurement of maturity when implementing threat hunting

Through combining CMMI, HMM and the SWOT analysis, we present a complement to the available measuring tools, with which an organization can determine if implementing threat hunting would be beneficial. This model will tell if your organization's conditions are either insufficient, sufficient, fair, satisfactory or optimal for a future threat hunting investment. Where an organization stands in terms of centralized logging (SIEM, quantity and quality of data collected), host and network visibility, budget, incident response and processes are the main factors when ascertaining an organization's capability level. The amount of threats to your organization should also be considered when deciding on a threat hunting implementation.

Table 3: A measurement of maturity when implementing threat hunting

Identifier	Capability level	Description
0	Insufficient	There is no centralized logging, routine data collection, host and network visibility or incident response. Staff are limited and budget does not allow for threat hunting.
1	Sufficient	Some routine data collection and Incident response is in place. Can make use of Threat Intelligence to conduct searches. Network and host visibility exists. Budget can be stretched for staff to perform for minor threat hunts. Processes are mostly reactive
2	Fair	SIEM is deployed with a high level of routine data collection. Incident response is in place. Some processes are proactive. Network and host visibility exists. Budget allows for some threat hunts.
3	Satisfactory	Able to generate new data analysis procedures. SIEM is deployed with network and host visibility and a high or very high level of routine data collection. Incident response team on standby. Budget allows for threat hunting. Most processes are proactive.
4	Optimal	SIEM is deployed with network and host visibility and a high or very high level of routine data collection. Incident response team on standby. Budget allows for having designated threat hunters. Tools and underlying strategies are in place. Processes are proactive,measured, controlled and are being optimized.

7 Constructing a framework for threat hunting

With the definition in mind, this part of the thesis seeks to build a framework for threat hunting. The model will have a strategy for making hypotheses, to reduce the risk of the hunt failing due to poor planning. The hypothesis and approach will be based on your goal, though some exceptions will be discussed later in the chapter. When testing the hypothesis, or encountering an adversary, an iterative strategy will serve as a foundation for threat hunting operations. A good iterative strategy will not only hinder adversaries, but also educate the hunters, improving their skills.

7.1 Iterative strategies

There are no rules for threat hunting, but having a defined procedure can benefit the hunt and organization. Having steps to follow will ensure that vital tasks are not missed and that hunters will know the next course of action. This section will introduce some of the most popular iterative methods that can be applied to cyber security. Although using military strategies for cyber security might seem controversial, war is the oldest and most studied adversarial conflict. The goal here will be to find a method that enables threat hunters to either increase the speed of their own process or decrease the speed of the adversary's, both resulting in a state of control. Controlling the tempo, you gain relative superiority over the opponent, interrupting his procedure. An equally important goal is to create an underlying strategy for daily threat hunting operations. The following iterative strategies do not originate from IT-security, but have been used for physical scenarios such as war and also in factories for quality assurance. Although most of them can be modified to fit into a threat hunting strategy, sometimes new strategies can deviate from the actual original iterative strategy.

7.1.1 F3EAD

Find, Fix, Finish, Exploit, Analyze and Disseminate (F3EAD figure 7) is a military methodology invented by the US Special Operations Forces and intelligence teams during the war in Iraq and Afghanistan 2001 to target terrorist networks. The F3EAD combines operations and intelligence to form a six-step iterative cycle. The operational steps focus on finding, fixing and finishing the adversaries and their operation. The first step, find, involves locating the target. The second step, fix, is immobilizing it. Finishing the adversary's operation and gathering relevant data is the third step in the operational part of the F3EAD method.

The intelligence part is initiated by processing the information gathered during the operational phase and then exploiting it to reveal more details about the adversary. Every piece of data is then analyzed for relevance and what can be used against other targets. The final product is then disseminated amongst allied forces.

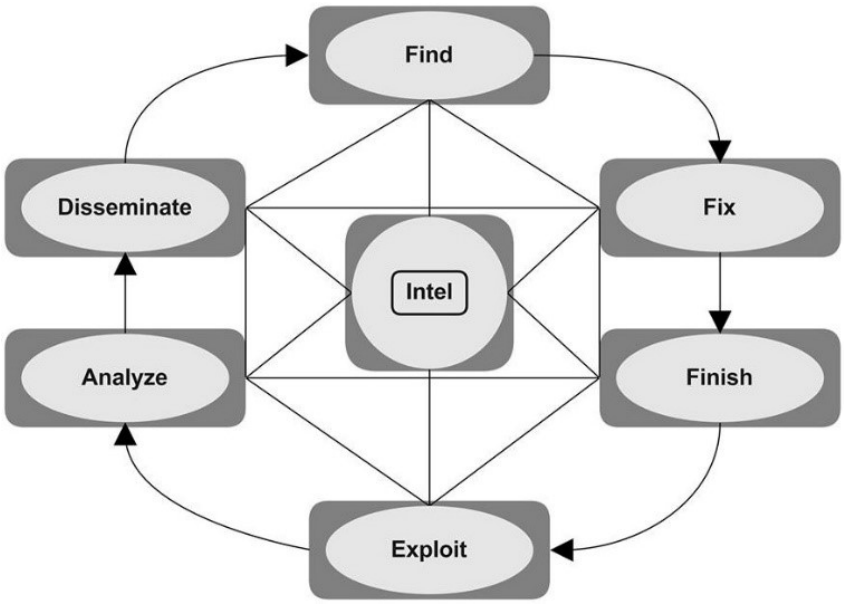

Figure 7: F3EAD loop

7.1.2 PDCA

Plan-do-check-act (figure 8) is a repeating four-step cycle invented by W. Edwards Deming who is considered the founder of modern quality control. It is an effective strategy for managing and continually improving processes and products. The plan part of the loop involves trying to understand the problem and coming up with a solution to implement. In the do-stage, the solution is implemented. In the next stage, the solution is observed in action. This stage involves checking if the solution works as expected, fixed or help to understand the problem, or if the solution can be improved in any way. The answer decides how to act; whether to permanently implement the new solution or to revert to the plan stage to revise.

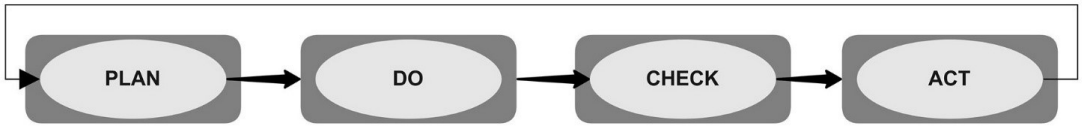

Figure 8: PDCA

7.1.3 OODA Loop

OODA stands for Observe, Orient, Decide and Act. The loop was invented by John Boyd, who was an air force pilot during the Korean War. Based on observations, dog fights and a pilot's decision making, he came up with the OODA loop (see figure 9). Harry Hillaker, who is the chief designer of the F-16, said about the OODA theory:

"Time is the dominant parameter. The pilot who goes through the OODA cycle in the shortest time prevails because his opponent is caught responding to situations that have already changed" [56].

The OODA loop is actually a natural way of reacting. An analogy could be a cyclist who needs to avoid an obstacle on his way to work. He observes a fallen branch in his way. He orients by considering his speed and the ability to stop or steer away. He decides to steer left and acts by doing so. Then the loop circles back to the observation stage until his next encounter. John Boyd wanted the OODA loop to be an explicit representation of the process that human beings and organizations use to learn, grow and thrive in a rapidly changing environment. Towards the end of his life, John Boyd drew a complex diagram to capture his vision of the OODA loop as a meta-paradigm for intellectual growth and evolution in a constantly shifting and uncertain environment [41].

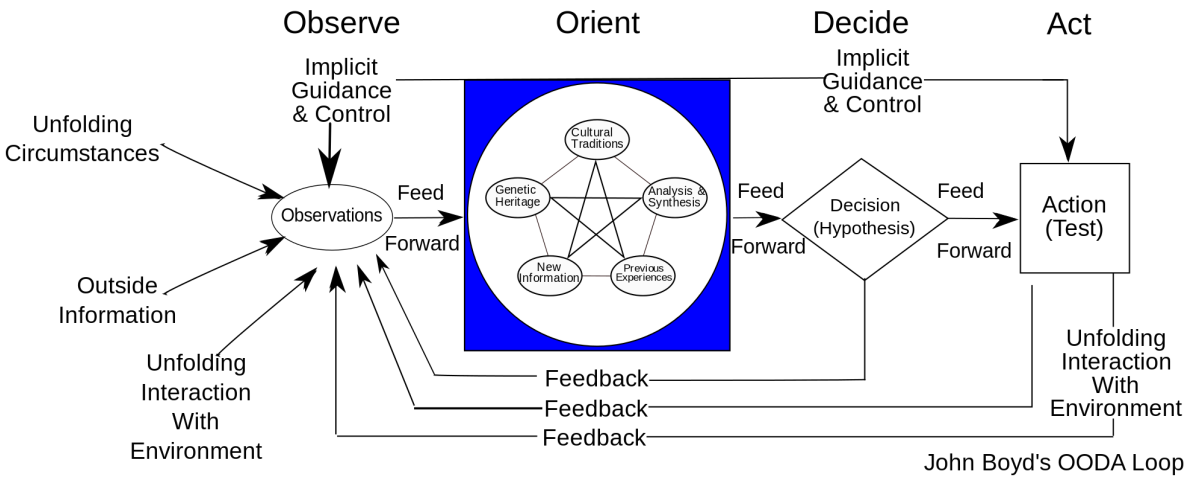

Figure 9: The OODA-Loop (Image created by: Patrick Edwin Moran)

Despite its name, certain scenarios make the OODA loop non-sequential. As illustrated above, all phases of the OODA loop can connect back to the observation phase, thus interrupting the sequence and making the process restart before the loop finishes. To use the aforementioned bicycle analogy, the cyclist could, after deciding to steer away from the fallen branch, observe another obstacle which he would collide with after acting on his decision to steer away. This would probably cause him to either steer in another

direction or stop altogether. Conclusively, the cyclist went directly from deciding to observing before acting and initially did not finish the loop. This is an example of the flexibility of a non-linear iterative strategy, where actions or goals can be changed on the fly.

Even though it is more a mindset than a strategy, the OODA loop is relevant in cyber security as it enables continuous improvement and offers flexibility. As a parable to dog fighting in the sky, today we may have an attacker versus a defender in an IT environment, and whoever goes through the loop faster will have the upper hand.

8 The Framework

The OODA loop will be used as the method for testing hypotheses and handling adversaries in this framework. This is mostly due to its adaptability, learning capabilities and non-linearity. F3EAD, although a viable strategy, might be too fixed on one specific goal and gives little or no chance to change requirements depending on the situation. In that sense, PDCA is similar to F3EAD.

The OODA loop does not require the certainty of a goal. Instead, it allows for changes during the hunting process and focuses more on introducing more intelligence to the hunter via feedback loops. Hunting for existing intrusions means trying to predict the adversary's course of action in an attempt to interrupt his cyber kill chain. Although PDCA also includes an observation or planning stage separate from the acting stage, something that offers an adaptive learning process, the problem with PDCA is its inflexibility; it relies more on the established sequence than the actual process of a live hunt [23].

8.1 OODA Loop details

8.1.1 Foundation of theories

Instead of publishing white papers on his OODA model, Boyd preferred conveying his theories on warfare through extensive briefings using 200-slide presentations. His audience mostly consisted of military personnel and powerful politicians, and although included, OODA was merely one of many ideas on grand strategy. Eventually, the briefings' content changed, resulting in there presently being no conventional OODA material available. Nevertheless, the OODA model should not be discarded as its intentional use is warfare and because it has obviously been evaluated and refereed numerous times.

In the paper "Destruction & Creation" [5], Boyd provides a philosophical groundwork for his theories regarding warfare. He assimilates Gödel's incompleteness theorem, Heisenberg's uncertainty principle, and the second law of thermodynamics to strengthen the development of the OODA Loop.

Boyd inferred the following from above mentioned theories in his paper:

- Incompleteness Theorem: Any logical model of reality is incomplete (and possibly inconsistent) and must be continuously refined/adapted in the face of new observations.

- Uncertainty Principle: There is a limit on our ability to observe reality with precision.
- Second law of Thermodynamics: The entropy of any closed system always tends to increase, and thus the nature of any given system is continuously changing even as efforts are directed toward maintaining its original form.

Having these considerations in mind, Boyd concluded that in order to maintain an accurate or effective grasp of reality, one must undergo a continuous cycle of interaction with the environment, prepared to adjust to constant changes. This was natural to Boyd and he thought he could not be first with this theory. He went on to expand on Darwin's Theory of evolution, suggesting that natural selection applies to social context, such as war or business, and is not only biological. Putting these two concepts together, Boyd stated that his decision cycle was the central process of adaptation and that increasing one's own rate and accuracy of assessment versus one's adversary's accuracy and assessment would provide a tremendous advantage in any form of competition. The key? Being able to adapt to change.

8.1.2 Processes of OODA

In Boyd's briefing he presented OODA as model of four processes interacting with the environment [23]:

Observe
To observe is to absorb information about your surroundings, either through the methods of perception, interaction or via a third party. The observation phase is offered guidance from orientation to compartmentalize your observation, but also receives feedback from the two last phases.

Orient
Also known as situation assessment or situation analysis, John Boyd goes to describe Orient as follows: *"Orientation, seen as a result, represents images, views, or impressions of the world ... Orientation is an interactive process of many-sided implicit cross-referencing projections, empathies, correlations, and rejections that is shaped by and shapes the interplay of genetic heritage, cultural tradition, previous experiences, and unfolding circumstances. ... Orientation is the schwerpunkt. It shapes the way ... we observe, the way we decide, the way we act"* [6]. Basically, the orientation phase affects all other phases and has a wide range of circumstances all taking part in shaping our perception of the world. Only after having oriented ourselves do we possess the ability to come to a decision.

Decide

Through orientation, hypotheses of the environment are formed. Choosing the hypothesis mostly in correspondence with reality and concluding whether or not to take action is the process of decision. The decision phase also provides feedback to the observation phase [23].

Act

Only physically executing the decided course of action is considered the act phase. The previously chosen method is carried out, resulting in a change of the environment and thus relocating us back to the observation stage. The act phase is offered guidance from the orientation as well as the decision phase [23]. Boyd puts emphasis on the unique and crucial feature of the OODA loop: tempo or cycle time. He concluded that in order to prevail, the speed of our OODA loop must exceed the speed of our adversary's. A superior scenario is acting to interrupt the adversary's OODA loop [6].

8.1.3 OODA as an iterative strategy for the framework

Finding existing intrusions is generally reactive. But finding the possibility of an intrusion and offering these to a feedback loop like OODA can be considered proactive. These types of strategies are found in sophisticated SOC or security teams, where you increase your tempo to surpass the attacker's, gaining relative superiority in order to prevent, block, recover and learn via feedback.

8.1.4 OODA Usage and potential

The progression of the OODA loop is not absolute, putting more emphasis on individual capabilities. Having people perceive things in a holistic manner may prolong the reaction or even the proactivity of the operation. The optimal use of the OODA loop is not individually but more as a concept for the whole team. This offers the possibility of voicing different viewpoints as well as being able to consider all aspects. The outcome of improving your own speed as opposed to reducing an opponent's might be the same. Therefore, the tempo or agility of the hunt, something relative capable of changing over time, should be of greater importance than speed alone [27].

OODA is used in red teaming (stealth/ghost level penetration testing) [16], which is related to threat hunting and incident response. OODA can be used in a linear threat hunting campaign: you create a hypothesis, search, find, respond and learn. The orientation in the OODA loop is a starting point, and the observation phase is used for feedback during operations.

OODA can be put to good use in situations of stress and disorientation. As stress can lower performance, having to think in terms of F3EAD, IR-process and PDCA could potentially diminish operations further. OODA can be used both iteratively and non-iteratively, and can also collaborate with any other hunting method.

8.1.5 Working and learning with OODA

In the paper "Unifying Planning and Control using an OODA-based Architecture" [23] the authors claim that OODA lacks a learning process.

Most people think of OODA as a method solemnly for making decisions. However, the usage of OODA reaches far beyond that [28]. Overlooking the learning capabilities is common when looking at OODA as a linear PDCA-type loop. OODA has feedback loops, which essentially increases your knowledge and is, in other words, learning. Using OODA as a non-sequential but iterative method would provide a way of both working and learning. A threat hunting team could start with little or no information. Using OODA, the team would orient itself, then observe, then orient based on observations, leading to deciding and acting. Using the information for new observations and as feedback loops back, you construct knowledge as the team becomes more of an entity.

When constructing a team, one should consider that the orientation phase is individual. Recognition-primed decision (RPD) [33] defines how experienced people make quick decisions under time pressure. In this model, one decision-maker produces a potential course of action, compares it to circumstantial limitations and finally selects the first one that is unchallenged. RPD has been observed in various professions including nurses, fire fighters, chess players and stock market traders. The reason threat hunting benefits from this is because using RPD provides fast decision-making in stressful situations with limited information [64]. For optimal results, RPD requires experienced and skilled people. However, the orientation phase of OODA can still be manned by less experienced people as long as skilled people guide the decision-making. This allows for the less experienced to learn. Compared to F3EAD, OODA has fewer steps. Limiting the amount of methods is potentially better in stressful situations, especially with inexperienced personnel.

8.1.6 OODA and threat hunting

For threat hunting, the hypothesis formed in the orientation phase will be the focal point. Even though orientation is the second phase of OODA, threat hunting operations use the hypothesis in a proactive search, observing in all directions. In a scenario where the hypothesis is in line with the goal to find a way to breach a system, you start to iteratively build a situational awareness as you learn and discover more traits until a weakness is found. Results should be reconstructed in the orientation process as you go, trying other experiments in order to gain more information to feed back to the observation phase. This will result in the ability to go through the loop faster, gaining control of the tempo and an intuitive orientation and decision. Even though a goal should be in correlation with the hypothesis, you initially do not need to know what the goal will look like when achieved. Depending on the goal, hunting can merely involve proving the hypothesis. An example of this could be when you suspect you are breached but you need to find evidence of it.

8.1.7 Advanced persistent defense

We can see examples of the OODA loop being used in Sqrrl's Framework for Cyber threat hunting [57]. David Bianco, security architect at Sqrrl, has developed a model for threat hunting inspired by John Boyd's OODA loop. He calls it the Advanced Persistent Defense (APD, illustrated in figure 10) [59].

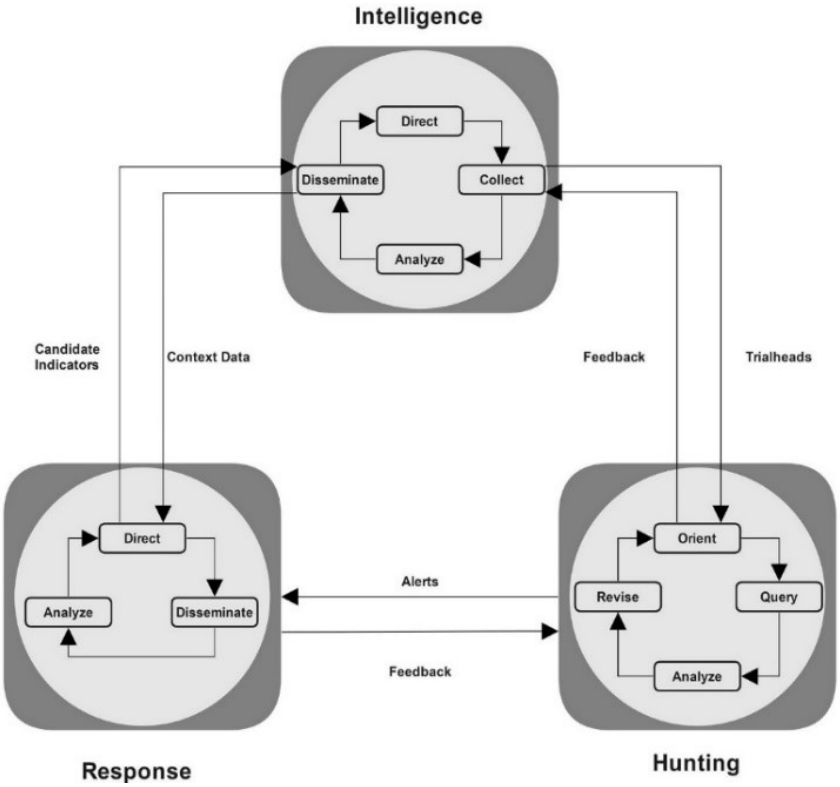

Figure 10: Advanced Persistent Defense (Recreated from A Framework for Cyber Threat Hunting Part 2: Advanced Persistent Defense)

The APD is made up out of three cycles that feed each other: Intelligence, Hunting and Response. The intelligence cycle internally consists of four steps. This first cycle aims to create an awareness of current threats, vulnerabilities and assets in your environment. Direct the construction of defenses by identifying assets and vulnerabilities. Know what intelligence you want to collect. Collect data from relevant sources. Analyze your data with automated tools. Disseminate the information gained from the previous step, and let that influence the construction of your defenses.

The hunting cycle aims to find APT's by going through data in a proactive and iterative manner. Orient the direction of your hunt based on hypothesis, IOC's and algorithm results. Query the information needed to reconstruct the progression and context of an attack. Analyze patterns and anomalies and compare to normal conditions. Identify signals of the adversaries' kill chain. Revise the queries to filter out less useful information.

The response cycle is the final circle, focused on mitigating the consequences of an incident. Contain the threat by restricting its access. Investigate the damage done by developing a narrative of the attack. This will be fed to the hunting cycle for digestion. Remediate the effects of the incident back to the intelligence cycle.

In the intelligence cycle, hypotheses, IOC's and algorithm results are created and subsequently fed to the hunting cycle. There, it is investigated and irrelevant data such as false positives are discarded. Threats found will be fed to the response cycle where they can be taken care of. Findings are fed back into the intelligence cycle and the process starts over. The APD model evolves over time as analysts gather more information on attacks. Although this method works in many cases, there is still a problem with the model being closer to linear. The point of OODA is the ability to change the direction instead of having to wait for the full APD-type loop to turn. The information gathered along the way cannot be implemented immediately, something which could have caused a pivotal change of the situation. Running multiple APD loops could be challenging, due to its complexity.

8.2 Threat hunting approaches for the framework

In the paper "*Diamond Model of Intrusion Analysis*", Sergio Caltagirone, Andrew Pendergast and Christopher Betz introduce an unprecedented way of analyzing intrusions [8]. The diamond model (illustrated in figure 11) is composed of four core elements: adversary, infrastructure, capability and victim. According to the authors, all malicious activity contains the features found in the edges of the diamond. Naturally, any threat hunting approach contains a combination of these components.

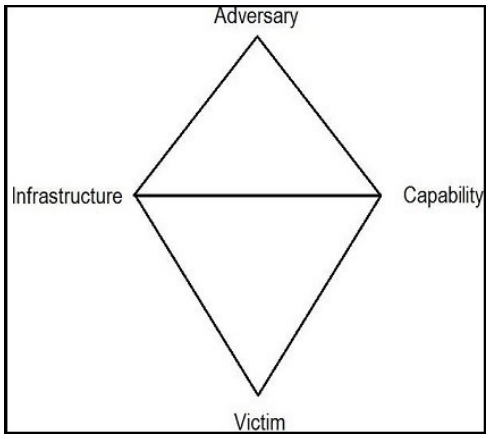

Figure 11: The diamond model (Recreated from Diamond Model of Intrusion Analysis)

Prior to initiating a hunt, it is beneficial to have a strategy, since threat hunting is both costly and difficult [9]. Hunting is also a risk since resources will be spent on an endeavor that in some cases does not yield any results. Having a strategy will significantly increase chances of success and a positive return of investment. Sergio Caltagirone [9] has composed four questions, the answers to which are crucial when forming hypotheses and strategies for threat hunting:

1. What are you hunting? Define exactly the point of the hunt. Are you searching for exploration, lateral movement, exfiltration et cetera?

2. Where will you find it? What you are searching for regulates your point of inspection. Specify which information to examine to find the given type of malicious activity.

3. How will you find it? Not all hunts require the same tools, which means that to answer this question you must decide which tools will be most effective.

4. When will you find it? To avoid spending an excessive amount of time on a specific hunt, create a time frame which you are supposed to dedicate to the hunt. If unsuccessful in finding anything, you terminate it.

Taking these questions and answers into consideration when making a hypothesis for a hunt will increase chances of success. An additional means of success is creating many hypotheses at once, since each hypothesis can lead to a failed hunt. Having more redundant ways of approaching things generally increases your chances of success and also saves time. This is where the OODA loop proves especially useful; being able to change and add hypotheses instantaneously because of the feedback loops.

The creators of the diamond model made it non specific by design, allowing for expansion and flexibility. The intention was to simplify an inclusion of new theories [8]. The model presents several approaches that can be adapted to threat hunting. For the threat hunting model of this thesis, the four different approaches of the diamond model will be used for creating hypotheses, and the OODA loop will be the underlying strategy for testing the hypotheses.

8.2.1 The capability-centered approach

Employing the capability-centered approach involves analyzing the capabilities used against you by the adversary. The investigation may provide the organization with information regarding potential victims, infrastructure and technology supporting the capability and sometimes even clues about the adversary.

A prime example would be an analysis by Symantec and CrySyS where a link between Stuxnet and Duqu was uncovered. Upon reviewing the code, they discovered significant similarities suggesting they had been written by the same author. The complexity of the features used were so unusual it allowed for them to narrow down the list of potential adversaries [8].

Examples of hypotheses in a Capability-Centered Approach [9]:

1. *We hypothesize that network defenders share adversary capabilities via VirusTotal.*
2. *We hypothesize that we can identify unique malware via a malware zoo using static analysis.*

8.2.2 The victim-centered approach

In a victim-centered approach, the target of the adversaries is the central component for the hunter and being referred to as the victim. Public announcements made by adversaries to combine their efforts and target a single person is an example of an offensive victim-centered approach. Most network defenders would likely respond with a defensive victim-centered approach as it provides a high chance of revealing adversaries. Additionally, when data related to the victim is being examined, it can result in discovering other elements of the diamond model related to the adversaries.

The Honeynet Project is an exceptional example of a victim-centered approach; using a specifically configured host as bait, adversaries were invited to exploit it, thus revealing their infrastructure and capabilities. This revelation supplies threat hunters with valuable information that can be used for mitigation or future reference. The only problem with a victim-centered approach is that an overload of malicious activity may eventually exhaust inexperienced threat hunters.

Examples of hypotheses in a victim-centered approach [9]:

1. *We hypothesize that several adversaries target a specific victim within our organization*
2. *We further hypothesize that these adversaries deliver their payloads through email as attachments or hyperlinks.*
3. *Our hypothesis is strengthened through analysis of machine data from email servers collected by our SIEM.*

These assumptions answer three of the four previously mentioned questions: the what, where and how. The next step could be to investigate the findings by detonating an adversary's payloads in a controlled environment to reveal the tools and techniques the adversary is using. This information would serve as feedback in the OODA loop for mitigation and learning. To answer the fourth question of when, a specific amount of time should be set for the task of detonating payloads.

8.2.3 The infrastructure-centered approach

When using the infrastructure-centered approach, you focus, unsurprisingly, on the infrastructure of the adversary. Through studying this, you may gain additional information about related elements: you might uncover other victims connected to the same infrastructure or you may gain clues from spotting capabilities used by the infrastructure. Discovering new infrastructure could possibly change the direction of the hunt, and should be brought to attention via the feedback loops of the OODA loop.

A concrete example of an infrastructure-centered approach was done by the Command Five team during their SKHack investigation [20]. After analyzing malware used in the attack, the team discovered multiple callback domains which they resolved to IP addresses. Subsequently, the IP addresses were entered into the WHOIS register where the team observed several other domains with the same registrant. Although the other domains had not been used in the attack, the team had reason to believe they were controlled by the same adversary. This led to proactive countermeasures, hopefully preventing or at least postponing future attacks by this adversary [8].

Examples of hypotheses in an Infrastructure-Centered Approach [9]:

1. *We hypothesize that adversaries have established infrastructure prior to attacks*
2. *We hypothesize that adversary X keep structuring their domains using the pattern badthings_ <victimname>.com*
3. *We hypothesize that adversary X continues to use the name server evildomain.com to host their infrastructure.*
4. *We hypothesize that monitoring the name server for a week for new names matching the pattern, we would find new names prior to their attack.*

In this scenario, the hunter could monitor the name server in a given time period by querying it daily for all domains in order to identify the domains not seen previously. Again, new observations are then fed back into the orientation phase in the OODA loop.

8.2.4 The adversary-centered approach

Probably the most difficult method is the adversary-centered approach. Utilizing it, you surveil all the actions of the adversary to gain insight to their capabilities and infrastructure. Although this could generate very precise and useful information, the obvious problem is that it requires constant access to the adversary.

As an example, a hacking group called the "Phonemasters" were monitored by the FBI. The information gained from their phone calls and modem activity allowed the FBI to unveil their entire operations including all other elements: their victims, capabilities and infrastructure [8].

Examples of hypotheses in an Adversary-Centered Approach [9]:

1. *We hypothesize that adversaries are using personas to register malicious domain names.*
2. *We hypothesize that some of these domain registrations aliases could relate to real people.*
3. *We further hypothesize that adversaries have mistakenly tied their operational alias to their real personas revealing their personal details.*

8.3 Gathering intelligence for the hunt

Data is essential when doing a threat hunt. It is what the analysis revolves around. The following are suggestions of environments where a hunter could start when making a hypothesis and picking an approach.

- **Clients**
  Operating systems, Dump data, Application specific logs, Security applications.
- **Networks**
  IDS/IPS, Firewalls, Web proxies, VPN End Points, Centralized authentication, Routing/Switching infrastructure, WLAN infrastructure, Antivirus, Proxies, Honeypots/Spam traps, DHCP logs.
- **Internet**
  Forums, Mailing lists, Blogs, Social media, Public/Private channels
- **Groups**
  Employees, Partners, Security researchers, Intelligence vendors, Government, Software vendors, ISP, Competitors.
- **Messages**
  E-mail/instant messaging (content, attachments, meta, source/destination), Voice mail, Call records

8.4 Spreading the findings of a hunt

The spreading of awareness of points of compromise and mitigation tactics is an important aspect of threat hunting in any organization. A successful threat hunt can yield various results. It can lead to finding a threat, exposing a weakness in the current security infrastructure or discovering a breach, sometimes even an ongoing attack. The hunt's primary objective is to mitigate all the points of compromise it reveals. Once a point of compromise is discovered, awareness of it must be communicated throughout the organization and, if applicable, a method of mitigation should be implemented in the entire organization. If the results of a successful hunt are not communicated, the hunting program will be inefficient and waste resources. In an organization with multiple large networks around the world, the spreading of information is not hard to achieve. Therefore, if handled correctly, the implementation of mitigation will likely scale well.

As a scalable model of spreading security information within an organization, we propose the use of a tree-like structure. The most basic example of such a structure can be represented as a binary tree(figure 12). A binary tree is a set of nodes (in this case represented by people and/or groups of people within the organization) where every node is a binary tree in itself. The root node of every subsequent binary tree is responsible for spreading information to its subnodes. Such a practice ensures that information reaches every part of the tree and in case it fails, it is easy to find the broken link in the chain of information propagation.

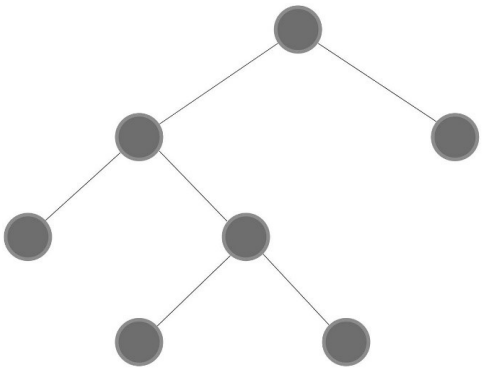

Figure 12: Binary tree

What we propose is to create this structure in the following manner:

The hunters are the root of the whole tree and are responsible for initiating the spread of information in the organization. They forward their findings to all their subnodes – for example heads of different departments. These subnodes are in turn responsible for forwarding the information to all their subnodes (for example project leaders in departments), which again are responsible for their subnodes and so on. This ensures that each node in every subtree receives the necessary information. Moreover, control of information spread in such a structure is easily obtainable - either by asking the lowest hanging nodes if they received the information, or by deducing the broken chain of spread information based on behavior that contradicts given instructions. If a node has acquired insufficient or incorrect information, one checks its root node and continues upwards to find the node that has received information without forwarding it properly.

A binary tree can be upgraded with a more advanced structure so that the number of subnodes per node is tailored to the needs of the organization, but this most basic tree shows how scalable a binary tree is. The number of nodes is practically limitless and it works well for both small and big organizations. Because each subnode is also a binary tree in itself, one can manipulate the structure by merging or dividing such subtrees from the core structure without affecting other trees. The amount of work per node is spread in a wide spectrum, ensuring that the amount of information transferred per

node is minimal and contributes to a system less prone to "bottlenecking". The number of steps to deliver a message from the root (hunters) to its farthest subnode can be easily calculated by defining the depth of the tree; For example an organization with 100 employees structured in a binary tree T would have a depth of:

$$D(T) \geq log_2(100) \approx 6.64$$

For a message to reach all 100 employees it would have to go through maximally seven nodes.

8.5 Example of a hunt utilizing the framework

Following our threat hunting framework, we begin with answering the four questions of what, where, how and when for our hypotheses, goal and approach. Several hypotheses are created and the analysis begins using relevant tools and techniques with OODA as an underlying strategy. For educational purposes, every phase of the OODA loop will be pointed out after each action in the procedure is described. Feedback throughout the hunt is fed to the Observe process and forwarded to the Orient process on the fly.

8.5.1 Background

Phishing mails containing potentially malicious attachments (executables and macros) and hyperlinks are getting through the spam filter and arriving at the inbox of important people at an organization. It poses a threat as they might download the attachment or follow the hyperlink, resulting in infecting the client and/or network.

8.5.2 Four questions

1. *What are you hunting?*
   We are hunting malicious macros hidden in seemingly legitimate office documents attached to email.
2. *Where will you find it?*
   We will find these malicious macros in office documents in emails, sent to employees in the company. The Internet will be used to acquire information on the matter.
3. *How will you find it?*
   The Internet will be used to compare other network defenders' approaches to the problem. The documents might get trapped in antivirus, or reported by employees. Our SIEM system could also be modified to alert.
4. *When will you find it?*
   We will look for documents suspected of containing malicious macros during 9am – 2pm today.

8.5.3 Objective

To stop malware from entering the network via employees' emails by stopping the malicious macros from executing, thus possibly infecting the client/network.

8.5.4 Approach

Identifying and defending against the tools of an adversary is to identify his capabilities, which would have made this a capability-centered approach. But in this scenario, we have specific people at the organization being targeted, so this will be a victim-centered approach. When finding the malware, focus will revolve around these people and their mailboxes and not on the actual malware.

8.5.5 Hypotheses

1. We hypothesize that documents containing malicious macros are being sent to employees via phishing emails.
2. We further hypothesize that we can block malicious macros from running in our environment using Office or group policies or registry editing.

8.5.6 Procedure

Our hypotheses is the center of the hunt, the orientation process in terms of OODA. It affects the way we decide what to do and how to act. With this as a starting point, we will observe our surroundings. (Observe)

Using Internet-based searches, we find that Microsoft has some articles on macro-based malware hidden in Office documents. After speaking with members of the security team, we could get our hands on one of the Office documents through email picked up by Proofpoint Advanced Threat Protection. (Orient. Decide. Act. Observe. Orient)

Preventing all emails with attached Office documents from being sent in the organization is not an option since it is a vital way to share data. We only want to stop those containing malicious files. Since macros in Office documents calling to Internet sites are not common in legitimate Office documents, we will follow through with our hypothesis and try to see if there is a way to stop them from being executed. (Decide)

Upon opening the Word document containing the malicious macros we see that Word protects the user by opening any file that did not come from the user itself in protected view (figure 13). The protected view is a sandbox mode that does not allow for anything to run, but also prohibits the user from editing the document. (Act. Observe. Orient)

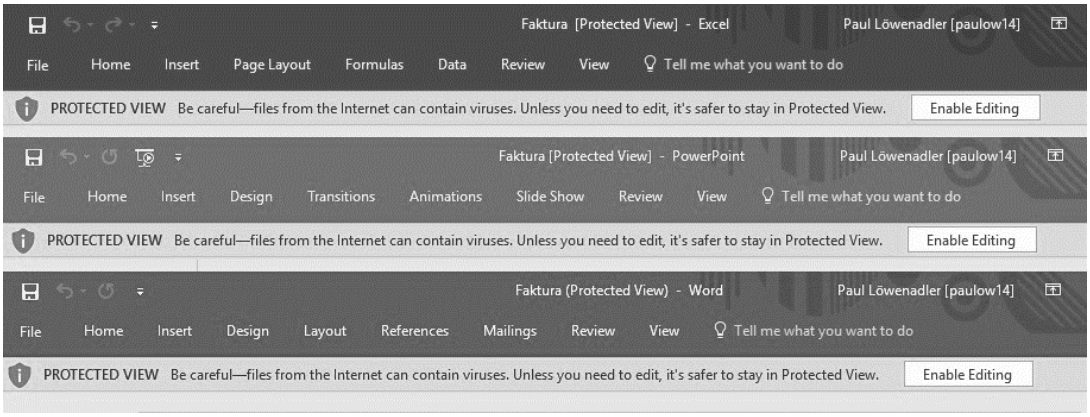

Figure 13: Protected view in MS Office programs

Although the protected view is doing its job and warning the user, it is not claiming that the file is malicious. It will look like this for all incoming office documents. Ultimately there is nothing stopping the user from removing the protected view, thus automatically running the malicious macro. Through the group policy editor (figure 14), we find a setting to block macros in Office files from the Internet. With administrative privileges, the following was done (Orient. Decide. Act):

```
Run > gpedit.msc
User configuration>Adminastrative templates >
Micosoft [Word/Powerpoint/Excel] 2016 >
[Word/Powerpoint/Excel] options > Security > Trust Center
```

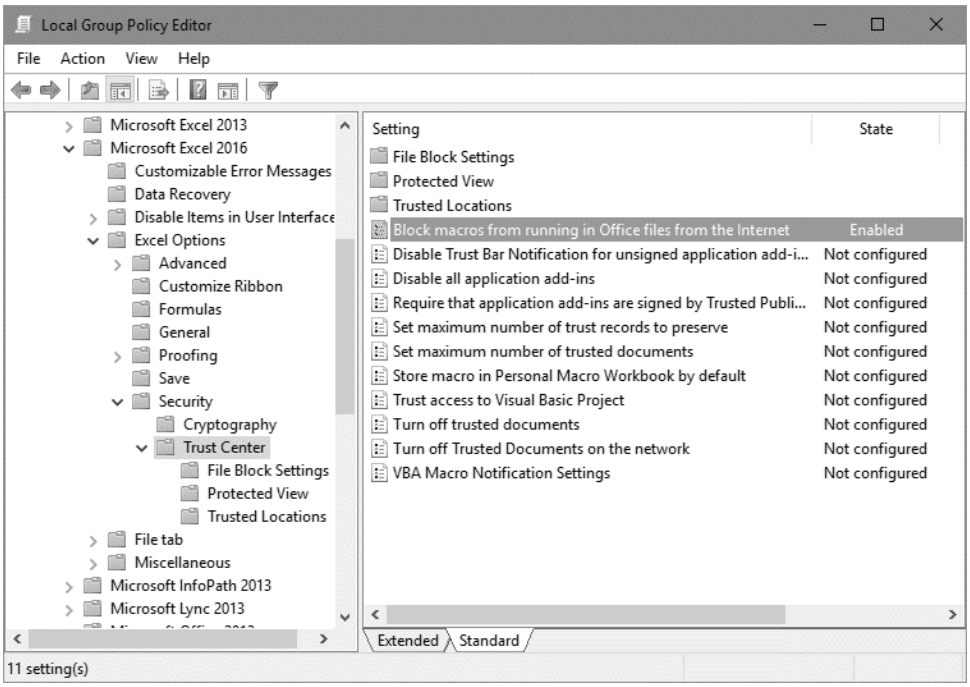

Figure 14: Local Group Policy Editor of a domain controller

We select the highlighted setting and enable it. This has to be done for every Microsoft Office program individually since there is no global option across the whole Microsoft Office suite. (Decide, Act, Observe, Orient)

8.5.7 Result

With the new setting blocking macros from running in Office documents from the Internet, this is what it will look like upon opening the document (figure 15). There is a red banner warning the user that this content is blocked and only an administrator can remove the settings.

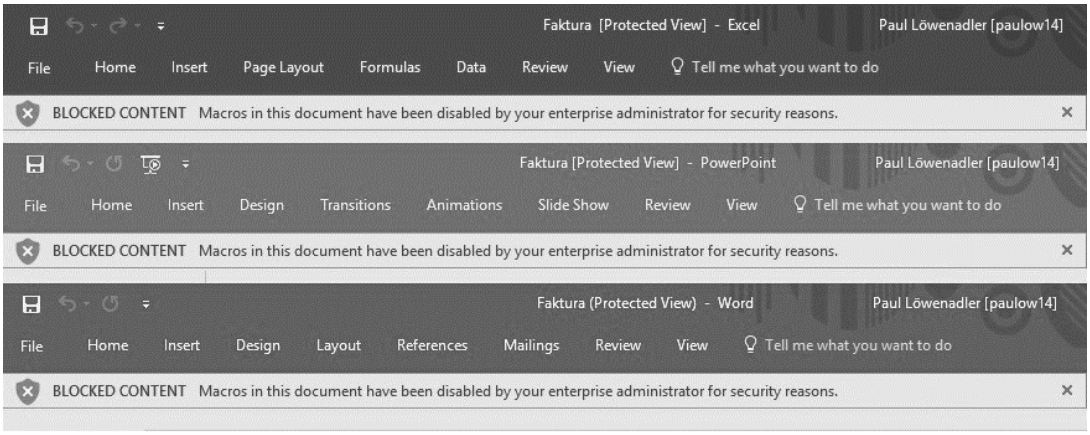

Figure 15: Blocked content in MS Office programs

8.5.8 Scaling the findings of this hunt

This hunt resulted in a mitigation technique to be implemented by an administrator of a domain controller. At the organization, there are ten domain controllers administered by five people. Furthermore, these domain controllers affect 120 000 client systems. Utilizing the binary tree structure to spread the information, the burden of work becomes minimal for the hunter, administrators and clients (illustrated in figure 16).

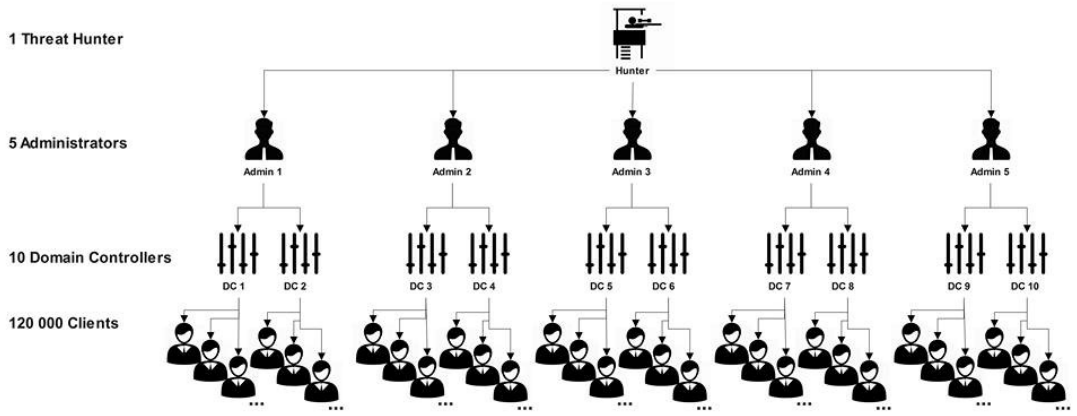

Figure 16: Organization in a binary trees

Here the message only has to go one step (from the hunter to the administrators) for it to affect over 100 000 clients.

8.6 Concluding the framework

1. Use the diamond model to pick an approach in correlation with the purpose/goal of the hunt
2. Create one or more hypotheses related to the approach and purpose/goal
3. Use OODA as an underlying strategy for the hunting process. The hypothesis is the Orientation of the cycle. Investigate using Tools and techniques.
4. Analyze results, recreate scenarios, conduct experiments. Information feeds back into the Loop.

9 Discussion

Results of this thesis provide the reader with a definition, a framework for conducting threat hunting and some tools to evaluate an organization's hunting potential. An example of a simple threat hunt is presented to show how our model works in practice. Furthermore, we show how results of these kind of hunts could scale and how threat hunting can be an addition to IT security operations at an organization.

Since there are businesses that revolve around threat hunting, there is currently an interest to keep it marketed and seemingly a new, amazing thing that should be considered by every IT company, which it probably should be. When researching threat hunting to gather material for this thesis, there seemed to be many papers released. However, we discovered they were all white papers, sponsored by the big vendors in the threat hunting business. When using white papers as a reference, one needs to consider the fact that they are not peer reviewed which means they can be released in virtually any state the authors see fit. We had to ask ourselves if there was any ulterior motive to their definitions and models of threat hunting.

Due to the lack of scientific papers on the matter, we felt the need to create a definition of threat hunting that was guaranteed to be impartial. We did this by researching the available material and then formulated one definition that neither contradicted nor plagiarized the other opinions. An alternative method, to backtrack the origins of threat hunting, proving where it came from and finally analyzing if it is a new strategy or not, would have been a project in itself.

Threat hunting is a diffuse phenomenon and depending on whom you ask, they might include or exclude certain procedures. We also noticed this while interviewing employees during our pre-study. A problem with defining a term as recent and dynamic as threat hunting is that even though our definition is sufficient for this thesis and is currently accurate, a year from now, our definition might be outdated.

As seen in the SANS surveys, many organizations benefit from threat hunting regardless of what their original idea of it is. However, organizations that get into threat hunting without having any plan or framework are most likely wasting resources to some extent [18]. Since threat hunting is related to other IT-security practices, a definition also serves as guidelines for the threat hunting procedure; if an incident response team is already in place, there is no need for a threat hunter to do their work. Although organizations are known to use their IR team as hunters, when scaling to the infrastructure of a big organization it would quickly get very expensive to involve the team for every hunting expedition. We have learned that the IR team at some organizations costs several thousand

Swedish crowns per hour, resulting in some incidents costing roughly a million. When the IR team is evaluating their performance, they can probably conclude that threat hunting around the targeted area could have resulted in avoiding the whole incident. In our opinion, threat hunting will have the least economic impact if carried out by individuals or smaller teams of hunters. An organization could either use existing IT personnel but ideally dedicate one or two employees specifically to threat hunting. However, sometimes employees with other IT duties might have to be temporarily allocated to threat hunting. This works well with OODA as an underlying strategy for threat hunting operations, as having an experienced hunter leading the orientation and decision phase allows for less experienced personnel to learn and contribute.

Before implementing threat hunting, some criteria should be met, as shown in the SWOT analysis. An organization suffering from shortage of security personnel with sufficient capabilities, and thus having issues with staying on top of sophisticated threats, could look into outsourcing threat hunting. Today, a handful of companies offer threat hunting as a service. These services generally consist of continuous endpoint monitoring, exposure, and elimination of threats, but also provide an organization with a to-do-list of vulnerabilities to prevent future attacks and lesser the burden of the incident response [14].

What we find interesting is the interdisciplinary studies about military and IT-security strategies, something that can also be found in other threat hunting articles. The theatre of the battlefield in IT security and war differ, but the adversaries nonetheless need to be dealt with and defended against. Soldiers do not really need to adapt to new weapon systems to the degree that IT security personnel must. That being said, we wanted to find a non-linear and easy-to-follow iterative strategy that can adapt to the changes of the situation. An example of this is Sqrrl's APD, where OODA is utilized but not to its full potential. They offer a framework based on OODA, but we get the impression that their model is too advanced and relies heavily on technology, limiting hunters [59]. We believe the strategy should be in place independent of the technology, or one might become a prisoner of technological limits. Our interpretation and implementation of the OODA loop is used in its standard form, surpassing the risks of the loop becoming linear due to modifications.

Our framework for threat hunting provides a structured and flexible way to perform threat hunting. By incorporating the diamond model to generate good hypotheses and approaches, it creates distinct guidelines for the hunt. Answering the four questions of what, where, when and how is important in order to get a clear picture of what is supposed to be done. As stated previously, knowing the exact procedure saves time and other resources. Our framework allows for having several hypotheses and enables hunters to change hypotheses on the fly. Since OODA works in a nonlinear fashion, it is even

possible to change the goal. This can be done by swapping or modifying hypotheses, or running multiple OODA loops.

From research and interviews we have learned that many think of threat hunting as an art form or something impulsive that should not be standardized; one could argue that having a procedure to follow for every hunt may become tedious and might cause hunters to act recklessly. We think that the framework alone shows the benefits of having a structured way of conducting threat hunting operations. Some say that threat hunting should never be standardized because it would benefit the adversary to know how an organization approaches IT security. We mean that simply knowing about our framework is not enough to exploit it. The framework does not go into detail of exactly how to do things and is also flexible due to the nature of OODA. If anything, this would deter the adversary more than attract. However, our description of the OODA-loop may be of use to adversaries though in a more general sense. Our framework is suitable for both general threat hunting operations and also for live adversary confrontation. The OODA loop has seen many days in combat, both physical and logical, and has proven to work [23].

Allotting a set amount of time to conduct the hunts would not only save resources but allow for a unit of measurement when investigating the opportunities to scale the threat hunting program. As with any project with a specified amount of time allocated to it, one can always divide the workload to finish faster or to get more accomplished in the same amount of time. However, because of the abstract nature of threat hunting, it is difficult to say whether an increase in resources is going to yield a better result. Also, more manpower sometimes introduces new problems such as coordination difficulties. Nevertheless, there are scenarios that will scale well. One example of this is the abovementioned spreading of mitigation techniques. If information is spread in an efficient way, it also saves resources since hunters do not need to do work already done by others. As seen in our example hunt in this thesis (Chapter 9), the results accomplished by one hunter can quickly affect 120 000 clients in a network.

There are many more variables to successful threat hunting than automating the processes. The Hunting Maturity Model made by the threat hunting company Sqrrl advocates threat hunting tools for analysis and data collection, and determining the hunting maturity level of your organization solely through this model can result in the threat hunting process being limited by technology. Still, this model will probably benefit most organizations with the adequate economical resources to acquire these tools. Sqrrl offers software that will most likely put your organization at the highest hunting maturity level (four). However, one needs be aware that this maturity model is made by a for-profit enterprise and created to mirror their product. The HMM model thoroughly describes the different levels of maturity and as with CMMI could, combined with other means,

more thoroughly identify and measure an organization's capability. We presented a model inspired by the HMM and with CMMI as a foundation. Identifying an organization's capability using our model should give an idea of the resources that need to be spent in order to get a solid threat hunting program running [57] [11].

Depending on their goals and resources, threat hunting is a viable strategy for many organizations, as shown in our maturity model for organizations looking to invest. Robert Richardson states that even though Information Security has become one of the major concerns of today's firms [47], a proactive security investment is difficult to sell. Even if a proactive security investment would lead to fewer incidents overall, authors Qian, Fang and Gonzales claim that proactive approaches to security investments can be hard to motivate, and that there exists a paradox: An investment with proactive intent will both lower the frequency and lesser the consequences of an incident, which will lead to the organization having a lower risk perception [45]. This could consequently lead to challenges when justifying an investment.

"Nobody ever gets credit for fixing problems that never happened" [45].

We realize that implementing a whole framework or defining threat hunting for an organization can be problematic. But even if organization do not choose to implement our framework or definition, parts of it can certainly be put to good use. For example, knowledge of the OODA loop used for threat hunting and learning could possibly benefit many organizations. Additionally, simply by using parts of the definition an organization could realize that it has competence enough to begin with threat hunting.

10 Summary

Because of the current popularity of the phenomena, organizations that take IT security seriously cannot ignore threat hunting. Threat hunting is channeled through the Internet via blog posts, articles, webinars, seminars, white papers and even recently at big security conferences. However, there is still no scientific paper on threat hunting; this might be the first one. Threat hunting is viewed as an important addition to IT security as the proactive aspect complements the weaknesses of the old reactive rule-based defenses. Together they form a stronger defense, necessary to keep up with today's adversaries.

Having limited scientific material meant there was no coined definition of threat hunting. Hence, the first part of the thesis had to define what is referred to as threat hunting. Before investing in threat hunting, tools like the SWOT analysis, HMM and our own capability model can be used to help assess if threat hunting is a viable investment. The framework created not only puts existing strategies for IT security to good use, but also shines new light on old war strategies. Together they form a framework and the power of the OODA loop as a strategy and learning tool is explained. Using the framework, a hunt is demonstrated, along with example of how the result can be implemented into a larger organization. Ideally, threat hunting is performed continuously and automatically to raise the overall security. But assigning people to perform threat hunting and working shifts can be hard to motivate. This thesis offers alternative solutions. If resources are a factor, and an organization cannot afford to have their IR team or designated personnel working proactively, our framework allows for smaller hunting parties. Having one senior threat hunter leading operations and borrowing personnel with competence needed for the scope of the hunt is enough, as OODA allows for learning simultaneously, provided that someone with experience is in charge.

10.1 Future work

It would be beneficial to keep the definition and framework up to date by editing, but also by including new information, such as an easier step-by-step template for conducting frequent threat hunts. There is huge potential in using Artificial Intelligence (AI) and machine learning for improving and automating threat hunts, searching for patterns and getting automations to behave humanely. We would like to see how implementing the OODA loop can benefit a threat hunting party over time by conducting experiments and keeping statistics. As more and more papers will be written on the subject, threat hunting will become more established and used, thus easier to study.

References

[1] S. Alonso. *Cyber Threat Hunting (1): Intro*. 2016. URL: `www.sqrrl.com` [Accessed: 13 Mar. 2018].

[2] A. Barriero. *How to choose a SIEM solution: An overview*. 2011. URL: `www.techrepublic.com` [Accessed: 13 Mar. 2018].

[3] Kuhn D. Wells-C. Armitage J. Clark G. Cusick K. Garcia S. Hanna M. Jones R. Malpass P. Minnich I. Pierson H. Powell T. Reichner A. Bate R. *A Systems Engineering Capability Maturity Model, Version 1.1*. Tech. rep. Software Engineering Institute, 1995.

[4] Waters S. Bell J. *Introduktion till forskningsmetodik*. fifth. Studentlitteratur AB, 2015.

[5] J. Boyd. "Destruction And Creation". MA thesis. US Army Command and General Staff College, 1976.

[6] J. Boyd. "Organic Design for Command and Control". Unpublished lecture notes. 1987.

[7] Grabosky P. Alazab-M. Chon S. Broadhurst R. "Organizations and Cyber crime: An Analysis of the Nature of Groups engaged in Cyber Crime". In: *International Journal of Cyber Criminology* (2014), pp. 1–20.

[8] Pendergast A. Betz-C. Caltagirone S. *The Diamond Model of Intrusion Analysis*. Tech. rep. Center for Cyber Intelligence Analysis and Threat Research Hanover MD, 2013.

[9] S. Caltagirone. *Building Threat Hunting Strategies with the Diamond Model*. 2016. URL: `www.activeresponse.org` [Accessed: 13 Mar. 2018].

[10] Millar T. Grance-T. Scarfone K. Cichonski P. *Computer Security Incident Handling Guide*. U.S Department of Commerce, 2012.

[11] Iacob I. Constantinescu R. "Capability Maturity Model Integration". In: *Journal of Applied Quantitative Methods* (2007), pp. 31–37.

[12] Hughes N. Davies M. "Doing a Successful Research Project: Using Qualitative or Quantitative Methods". In: *International Journal of Social Welfare* (2016), pp. 416–704.

[13] Ghemawat S. Dean J. *MapReduce: Simplified Data Processing on Large Clusters*. Tech. rep. Google, Inc., 2004.

[14] A. DeNisco Rayome. *Why threat hunting as-a-service is worth considering, but 'not a silver bullet'*. 2016. URL: `www.techrepublic.com` [Accessed: 13 Mar. 2018].

[15] M. Descombe. *Forskningshandboken : för småskaliga forskningsprojekt inom samhällsvetenskaperna*. third. Studentlitteratur AB, 2016.

[16] M. Devost. *10 Red Teaming Lessons Learned Over 20 Years*. 2015. URL: `www.redteamjournal.com` [Accessed: 13 Mar. 2018].

[17] Kokinov B. Leake-D. Turner R. Dey A. *Modeling and Using Context: 5th International and Interdisciplinary Conference, CONTEXT 2005, Paris, France, July 5-8, 2005, Proceedings*. Lecture Notes in Computer Science (Book 3554). Springer, 2005.

[18] E. Dr. Cole. *Threat Hunting: Open Season on the Adversary*. Tech. rep. SANS, 2016.

[19] G. Engel. *Deconstructing The Cyber Kill Chain*. 2014. URL: `www.darkreading.com` [Accessed: 13 Mar. 2018].

[20] Command Five. *SK Hack by an Advanced Persistent Threat*. Tech. rep. Command Five Pty Ltd, 2011.

[21] P. Gasper. *Cyber Threat to Critical Infrastructure 2010-2015*. Tech. rep. Idaho National Laboratory, 2008.

[22] Martin D. Nguyen-Duy J. Santana M. Schwartz E. Weber D. Girardi B. *Transforming SIEM Into An Early Warning System For Advanced Threats*. Tech. rep. RSA, The Security Division of EMC, 2012.

[23] T. Grant. "Unifying Planning and Control using an OODA-based Architecture". In: *Proceedings of the 2005 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries* (2005).

[24] D. Gross. *50 million compromised in Evernote hack*. 2013. URL: `www.cnn.com` [Accessed: 13 Mar. 2018].

[25] Shannon S. Hsieh H-F. "Three Approaches to Qualitative Content Analysis". In: *Sage Journals* (2009), pp. 1277–1288.

[26] M. Hurley. "For and from Cyberspace : Conceptualizing Cyber Intelligence, Surveillance, and Reconnaissance". In: *Air & Space Power Journal* (2012), pp. 12–33.

[27] J. Jaakonaho. *Information - it's about time*. 2016. URL: `www.linkedin.com` [Accessed: 13 Mar. 2018].

[28] J. Jaakonaho. *The unbearable lightness of a strategic thinking (also in Cyber Security)*. 2016. URL: `www.linkedin.com` [Accessed: 13 Mar. 2018].

[29] D. Jacobsen. *Vad, hur och varför? - Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*. first. Studentlitteratur AB, 2002.

[30] B. Johnson. *Threat Hunting as a Culture (HaaC): Moving Your Defences Towards an Aggressive, Proactive Style*. 2016. URL: www.carbonblack.com [Accessed: 13 Mar. 2018].

[31] Badger L. Waltermire D. Snyder J. Skorupka C. Johnson C. *Guide to Cyber Threat Information Sharing*. U.S Department of Commerce, 2016.

[32] J. Katz. *In the Matter of Jonathan G. Lebed*. 2000. URL: www.sec.gov [Accessed: 13 Mar. 2018].

[33] G. Klein. *Sources of Power: How People Make Decisions*. second. MIT Press, 1999.

[34] Carayon P. Duggan R. Kraemer S. "Red Team Performance for Improved Computer Security". In: *Sage Journals* (2004), pp. 1605–1609.

[35] Lee R. Lee R. M. *The Who, What, Where, When, Why and How of Effective Threat Hunting*. Tech. rep. SANS, 2016.

[36] J. Lejon. *Vad är Cyber Threat Hunting?* 2017. URL: www.linkedin.com [Accessed: 13 Mar. 2018].

[37] Guba E. Lincoln Y. *Naturalistic Inquiry*. SAGE Publications, Inc., 1985.

[38] N. Lord. *What is Threat Hunting? The Emerging Focus in Threat Detection*. 2016. URL: www.digitalguardian.com [Accessed: 13 Mar. 2018].

[39] MacAfee. *Operationalizing Threat Intelligence*. Tech. rep. McAfee, LLC, 2015.

[40] M. McIntyre. "The Integrated Product Development Capability Maturity Model (CMM)". In: *INCOSE International Symposium* (1996), pp. 574–576.

[41] McKay K. McKay B. *The Tao of Boyd: How to Master the OODA Loop*. 2014. URL: www.artofmanliness.com [Accessed: 13 Mar. 2018].

[42] A. Ommani. "Strengths, weaknesses, opportunities and threats (SWOT) analysis for farming system businesses management: Case of wheat farmers of Shadervan District, Shoushtar Township, Iran". In: *African Journal of Business Management* (2011), pp. 9448–9454.

[43] C. Osborne. *Most companies take over six months to detect data breaches*. 2015. URL: www.zdnet.com [Accessed: 13 Mar. 2018].

[44] Curtis B. Chrissis M. Weber C. Paulk M. *Capability Maturity Model for Software, Version 1.1*. Tech. rep. Software Engineering Institute, 1993.

[45] Fang Y. Gonzalez J. Qian Y. "Managing information security risks during new technology adoption". In: *Computers & Security* (2012), pp. 859–869.

[46] C. Rice. *Cyber Threat Intelligence*. Tech. rep. Payments UK, 2014.

[47] R. Richardson. *2008 CSI Computer Crime & Security Survey*. Tech. rep. CSI, 2008.

[48] Elgin B. Lawrence D. Matlack C. Riley M. *Missed Alarms and 40 Million Stolen Credit Card Numbers: How Target Blew It.* 2014. URL: www.bloomberg.com [Accessed: 13 Mar. 2018].

[49] S. Roberts. *Incident Response is Dead. . . Long Live Incident Response.* 2016. URL: www.sqrrl.com [Accessed: 13 Mar. 2018].

[50] S. Rosenblatt. *LivingSocial hacked; 50 million affected.* 2013. URL: www.cnet.com [Accessed: 13 Mar. 2018].

[51] M. Rouse. *advanced persistent threat (APT).* 2010. URL: www.techtarget.com [Accessed: 13 Mar. 2018].

[52] M. Rouse. *security information and event management (SIEM).* 2014. URL: www.techtarget.com [Accessed: 13 Mar. 2018].

[53] Mell P. Scarfone K. *Guide to Intrusion Detection and Prevention Systems (IDPS).* U.S Department of Commerce, 2007.

[54] S. Sorkin. *Large-Scale, Unstructured Data Retrieval and Analysis Using Splunk.* Tech. rep. Splunk, Inc., 2011.

[55] Baskerville R. Kim J. Spagnoletti P. "Incident-centered information security: Managing a strategic balance between prevention and response". In: *Information & Management* (2014), pp. 138–151.

[56] L. Spitzner. "The Honeynet Project: trapping the hackers". In: *IEEE Security & Privacy* (2003), pp. 15 –23.

[57] Sqrrl. *A Framework for Cyber Threat Hunting.* Tech. rep. Sqrrl Data, Inc., 2016.

[58] Sqrrl. *Cyber Threat Hunting.* URL: www.sqrrl.com [Accessed: 13 Mar. 2018].

[59] Sqrrl Team. *A Framework for Cyber Threat Hunting Part 2: Advanced Persistent Defense.* 2015. URL: www.sqrrl.com [Accessed: 13 Mar. 2018].

[60] P. Thomas. *Teen Hacker Faces Federal Charges.* 1998. URL: www.cnn.com [Accessed: 13 Mar. 2018].

[61] Various. *Strategy to Tactics: BIA for Actionable Insights on Adversary Behavior.* Online. Air University Behavioral Influences Analysis Center (BIAC) Workshop. 2007.

[62] J. Vijayan. *5 Things To Consider With A Threat Hunting Program.* 2016. URL: www.darkreading.com [Accessed: 13 Mar. 2018].

[63] J. Vijayan. *'Threat Hunting' On The Rise.* 2016. URL: www.darkreading.com [Accessed: 13 Mar. 2018].

[64] *Wikipedia:"Recognition primed decision".* Online, Wikipedia.

[65] *Wikipedia:"Security information and event management"*. Online, Wikipedia.

[66] *Wikipedia:"Trovärdiga källor"*. Online, Wikipedia.

paulow14@student.hh.se

thelil14@student.hh.se

PO Box 823, SE-301 18 Halmstad
Phone: +35 46 16 71 00
E-mail: registrator@hh.se
www.hh.se