TaHiTI: a threat hunting methodology

1 Introduction

Threat hunting is a relatively new area of expertise. While the activity itself is not new, specific hunting tools, models and best practices have been developed in recent years. As with any new area, there is often confusion on what exactly comprises this activity. Good definitions are lacking, as are common approaches on how to perform such an activity.

The 2017 SANS survey on threat hunting has indicated that only 4,6% of all companies engaging in threat hunting activities have adopted a published external methodology. Excluding outsourcing and companies that do not perform threat hunting, that leaves over 70% of organizations either using no methodology or a methodology that was created internally [[1]]. This shows a clear lack of availability of threat hunting methodologies that cover the entire process in a structured fashion.

Members of the Dutch financial sector that were conducting threat hunting activities have come to the same conclusion. In-house methodologies and hunting expertise were being developed separately. As such, the timing was right for a joint effort in creating a common understanding and common approach in threat hunting, as well as sharing best practices amongst each other. The **TaHiTI** (which stands for **Ta**rgeted **H**unting **i**ntegrating **T**hreat **I**ntelligence) methodology is a direct result of that effort. The methodology itself seeks to combine threat hunting and threat intelligence to provide a focused and risk-driven approach to threat hunting. Threat intelligence is used as a source for hunting investigations and is used throughout the investigation to further contextualize and enrich the hunt.

This article represents a brief version of the full **TaHiTI** documentation. The full documentation and supporting tool (MaGMa for threat hunting) can be obtained from:
https://www.betaalvereniging.nl/en/safety/tahiti/

2 Threat hunting

To have a common understanding of threat hunting, a common definition is required. Threat hunting in **TaHiTI** is defined as follows:
Threat hunting is the proactive effort of searching for signs of malicious activity in the IT infrastructure, both current and historical, that have evaded existing security defenses. This evasion of security defenses can be due to usage of new, improved or unknown attacker techniques, 0-day exploits or a lack of adequate detection technology within the organization. While incomplete or faulty configuration of detection technology or misinterpretation of

[1] https://www.sans.org/reading-room/whitepapers/analyst/hunter-strikes-back-2017-threat-hunting-survey-37760

security events by analysts during triage can be reasons for evasion as well, threat hunting assumes a properly running security monitoring process.

The main purpose of threat hunting is to reduce the time required to find traces of attackers that have already compromised the IT environment. By finding these traces as soon as possible, the impact of breaches to the organization can be minimized. Other benefits of threat hunting include:

- Identification of gaps in visibility necessary to detect and respond to a specific attacker TTP.
- Identification of gaps in detection.
- Development of new monitoring use cases and detection analytics.
- Uncovering new threats and TTPs that feedback to the threat intelligence process.
- Recommendations on new preventive measures.

2.1 Breach detection gap

As indicated, the goal of threat hunting is to decrease the gap between initial compromise by an attacker and the discovery of that attacker in the environment: the breach detection gap also known as *dwell time*. Figure 1 shows a timeline of an attack containing several key moments in time. The breach detection gap is the time between T=1 and T=2.

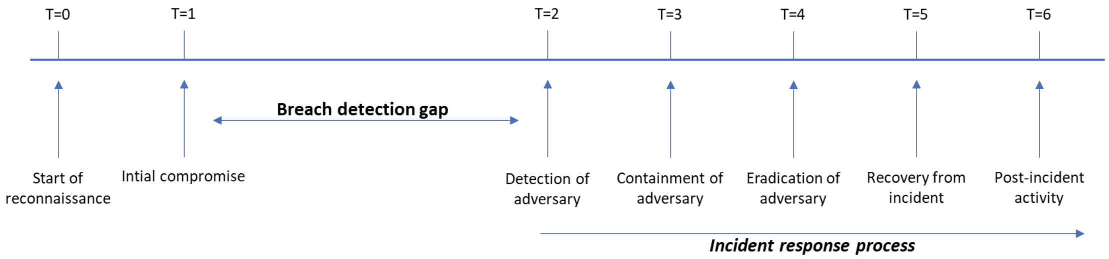

Figure 1: the breach detection gap

According to the latest Verizon Data Breach Investigation Report (DBIR), 68% of compromises went undetected for months [2]. Threat hunting plays an important role in reducing the breach detection gap. This is also evident in the SANS threat hunting survey, where improvements to incident response were mentioned as key improvements due to threat hunting activities. Threat hunting will aid to accelerate the detection of attackers by introducing new or improving existing detection mechanisms and thereby further closing the breach detection gap. Continuous insight into the state of detection mechanisms is required to avoid hunting for malicious activity that is already covered by traditional detection mechanisms. Use case management frameworks, such as MaGMa [3] can aid in such insight.

[2] https://www.verizonenterprise.com/resources/reports/rp_DBIR_2018_Report_execsummary_en_xg.pdf

[3] https://www.betaalvereniging.nl/en/safety/magma/

2.2 Types of threat hunting

When it comes to types of threat hunting, it basically drills down to 2 types: unstructured hunting and structured hunting.

Unstructured hunting is data-driven hunting. Potentially malicious activity can be detected by a hunter who is simply digging through available data looking for anomalies. This type of threat hunting does not start with a hypothesis, does not follow a predetermined path and is thus considered unstructured.

Structured hunting is hunting based on hypotheses: a hypothesis is created, the hunting activity is scoped and subsequently performed. **TaHiTI** is a structured hunting approach that involves several steps and a clear idea of what the hunters are looking for before any hunting activity is initiated.

2.3 Pyramid of Pain

The pyramid of pain [[4]] is an important and elegant concept that can be used in threat hunting and threat intelligence. The pyramid addresses how difficult it is for attackers to change certain characteristics of their attack. At the same time, it also shows how difficult it is for organizations to find these characteristics. Finding a file with a certain hash value is easy, but uncovering illegitimate use of PowerShell in an organization where PowerShell is commonly used poses an entirely different challenge. Similarly, it is trivial for attackers to generate a new file with a different hash, but much harder to move or modify an attacker technique to evade detection. Figure 2 shows the pyramid of pain.

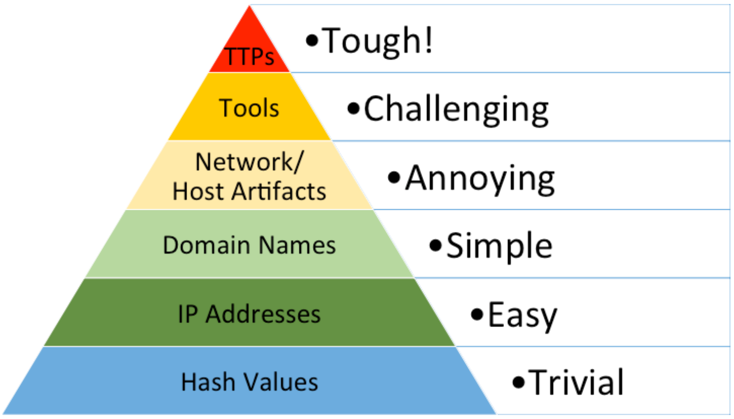

Figure 2: Pyramid of Pain (source: David Bianco, detect-respond blog). ***TaHiTI*** *focuses on the top 3 layers of the pyramid.*

[4] http://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html

The pyramid of pain connects threat hunting to threat intelligence. Threat intelligence provides relevant information on attackers on all layers of the pyramid. Threat hunting with the **TaHiTI** methodology will focus on the top 3 layers (but may use the lower 3 layers nonetheless) of the pyramid. TTPs (the methods of attack used by attackers) have the most focus in threat hunting investigations.

3 Threat intelligence

As with threat hunting, it helps to have a clear definition of threat intelligence. Threat intelligence is defined as follows:
Threat intelligence is the process of gathering, processing and dissemination of information about threats and attackers. The goal of threat intelligence is to contextualize the information and to deliver actionable information that can be used in the decision-making process.

The threat intelligence process puts information from the outside world into the organizational perspective and, if possible, advises on how to proceed. This requires determination of risk, impact and possibly mitigating measures from intelligence information. Threat intelligence also provides insight into how attackers operate, their motivation, the sectors and geographic locations they operate in and the level of capability they possess.

3.1 The relationship between threat hunting and threat intelligence

There is a clear relationship between threat hunting and threat intelligence. This has become apparent in the section on threat hunting, as some concepts in threat hunting are difficult to explain without basic knowledge of threat intelligence. For the **TaHiTI** methodology, 3 concrete elements of the relationship between threat intelligence and threat hunting are especially important:

- Intelligence as a starting point for hunting.
- Intelligence for contextualizing and driving the hunt.
- Hunting to generate intelligence.

3.1.1 Intelligence as a starting point for hunting

As threat intelligence provides us with a lot of information on attackers and their capabilities, it can be a major source for engaging in hunting activities. For example, a threat intelligence report describing an attacker group and their distinct capabilities should be of great interest. If that attacker group also operates in your organization’s sector and is also geographically relevant, the threat it poses may be significant. The threat intelligence process can trigger the threat hunting process based on this information and provide relevant context on the threat.

3.1.2 Intelligence for contextualizing and driving the hunt

During hunting investigations, threat intelligence can be used for contextualization of findings. For example, a certain TTP may be uncovered during the threat hunting process. Using threat intelligence, that information may be used to find related TTPs (for example, using the MITRE

ATT&CK framework [5][6]) or additional information on that TTP. This can subsequently be used to further drive the hunt. This process is called pivoting and may lead to additional hunting activities or refinement of the active hunt. For the **TaHiTI** methodology, this interaction between threat intelligence and threat hunting is especially important. Context from threat intelligence may lead to extending the scope of the hunt, adding new data to the hunt, refining the hunting hypothesis or generating ideas for subsequent hunts.

3.1.3 Hunting to generate threat intelligence

As mentioned earlier, threat hunting can be a source for threat intelligence. Hunting investigations may uncover previously unknown TTPs for attackers. This information can be used in the threat intelligence process to build an attacker profile. All such information can subsequently be shared with peers in threat intelligence communities, providing them with information regarding the uncovered threat. If these peers start their own hunting investigations based on this new TTP, they may uncover additional indicators that can be shared with the threat intelligence community. This way, a more complete picture of attacker capabilities and TTPs can be built in a community effort. Within an active threat intelligence ecosystem, the sum is greater than the whole of its parts.

4 TaHiTI methodology

As indicated in the introduction, **TaHiTI** stands for **Ta**rgeted **H**unting **i**ntegrating **T**hreat **I**ntelligence. *Targeted* because the methodology uses hypotheses to drive hunting activities. This means threat hunting is conducted with a specific goal in mind. *Integrating threat intelligence* because threat intelligence is a major source of threat hunting hypotheses, and is used to enrich and contextualize hunting activities. Lastly, threat intelligence may also be generated as a result of hunting activities.

4.1 The TaHiTI process overview

Figure 3 provides an overview of the **TaHiTI** process, and its 3 phases: Initiate, hunt and finalize. The process has 6 steps in total.

[5] https://attack.mitre.org/

[6] https://medium.com/mitre-attack/finding-related-att-ck-techniques-f1a4e8dfe2b6

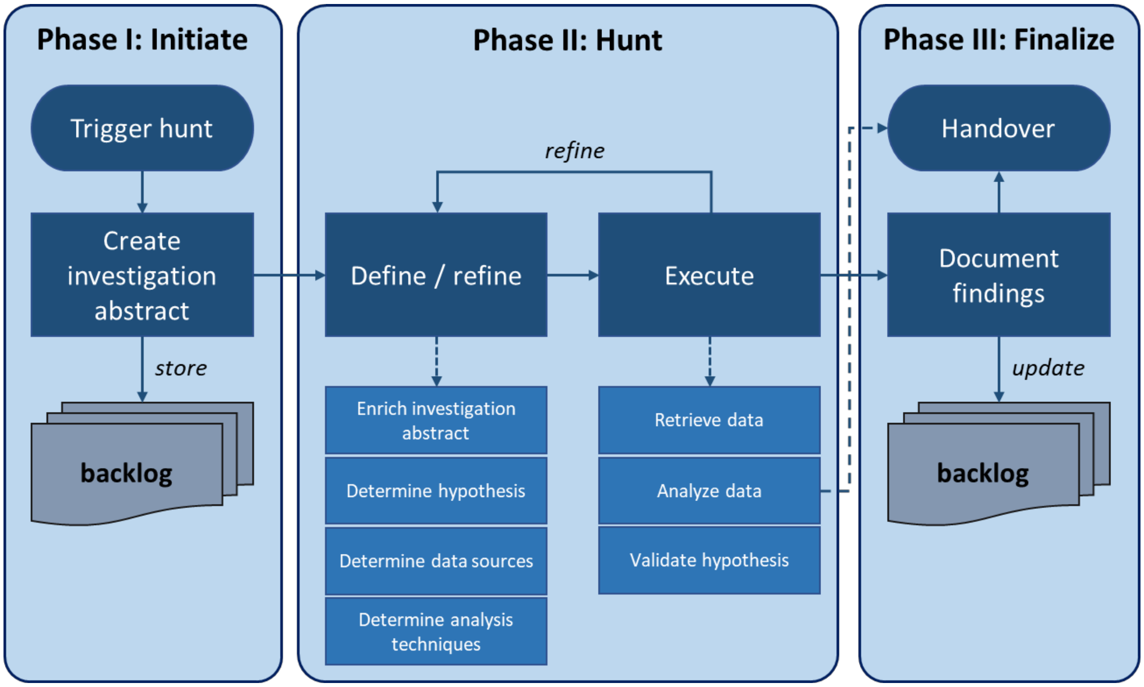

Figure 3: the ***TaHiTI*** *process*

4.2 Phase 1: Initiate

The initiation phase is where the input for threat hunting is processed. First, there is an initial trigger to initiate the hunting process. Next, the trigger is converted to an abstract of the hunting investigation and stored on the hunting backlog.

The threat hunting process can be triggered from several processes. Figure 4 shows triggers for threat hunting. An important thing to notice is that the processes that could potentially provide triggers to start the hunt strongly overlap with the processes that receive output from the investigation (figure 5). When executed well, hunting can act as an accelerator for improvement of these other processes.

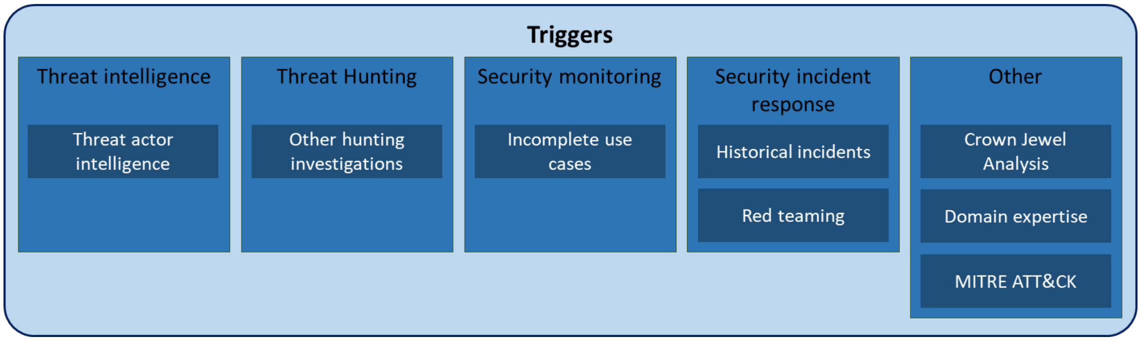

Figure 4: hunting triggers

Of all these triggers, the MITRE ATT&CK framework is especially important. The MITRE ATT&CK framework can be used as input for potential attack vectors and techniques, and contains a wealth of information for any hunter. The framework also provides suggestions for detection, which is valuable for both hunting and security monitoring. Note that this is not the primary purpose of the framework and should be treated as guidance for monitoring only.

When a trigger is received, the hunting team creates a hunting investigation abstract. This abstract does not include all details, but is a basic description of the investigation. Most of the information will be refined and updated in a later stage, when the hunt is selected for execution. After the abstract is created, it is stored on the hunting backlog. This backlog does not need to be a complex tool. Simple collaboration tools such as Microsoft SharePoint or JIRA can suffice. The most important thing is that the backlog provides the hunting team with the required insight to select the most relevant abstract for the next hunt.

4.3 Phase 2: hunt

The second phase of the hunt is where the actual investigation takes place. There are 2 activities in this phase. The first activity in the hunting phase is called 'define / refine'. The second activity is called 'execute' and is the actual conduction of the hunt.

In the 'define / refine' step, the details for the hunt are defined and made more concrete. During this step, the abstract is turned into an investigation by refining and adding information. Some new elements are added, such as required data sources and data analysis techniques. Most importantly, a hypothesis is created that drives the hunt. Generating this hypothesis is a vital step in the hunting process. A badly defined hypothesis will likely lead to no results or even worse, wrong results and thus wrong conclusions and recommendations to the organization.

With 'define / refine' activity completed, the 'execute' activity can be started. During the 'execute' step of the hunt phase, data is retrieved and analyzed. Existing hunting documentation lists a number of data analysis techniques. Some of these techniques, such as querying, are simple and easy to perform. Other techniques, such as clustering are more difficult to understand and require some basic understanding of statistics to use them properly. Some analysis techniques can be applied manually by analysts, while other require some form of machine learning. Hunting platforms that contain analysis techniques and visualizations can be leveraged to simplify analysis. In the data analysis step, the hunting team may find omissions introduced in the define stage. At this point, the hunters will refine the initial investigation. This is an iterative process that is repeated until the investigation is optimized. Refinements can be done to hypotheses, scope, selected data sources and analysis techniques.

When performing data analysis, threat intelligence can be used to add context to investigations. The need for this depends on the investigation. When the threat hunting team finds matches on specific TTPs, further analysis into that TTP must be performed. If possible, this activity should be conducted in collaboration with the threat intelligence team. Such analysis may provide information on possible threat actors, their methods and capabilities, technical infrastructure and other victims of the same actor (the 4 features of the diamond model of intrusion analysis [[7]]). The MITRE ATT&CK framework can be used this in this process. Additionally, the MITRE ATT&CK navigator, can be a useful resource as it associates attack techniques to APT groups. To determine which APT groups are relevant for your sector and organizational type, the APT threat tracking overview is a good starting point [[8]]. This information can subsequently be used to extend the hunting investigation to find additional malicious activity. This provides the hunter with a more complete overview of the compromise that has taken place.

The final activity of the 'hunt' phase is hypothesis validation. When the hunting investigation is finished, the hypothesis must be validated. This will either result in a proven hypothesis (malicious activity found, incident response started), disproven hypothesis (no malicious activity found) or inconclusive. In the latter case, the hunter can cycle back to the first step (define / refine) to change some of the parameters of the hunt and repeat the execution. In some cases, the required data may simply not be available. In such case, the hunt can be considered to be failed. This can still lead to valuable lessons learned.

4.4 Phase 3: finalize

The final phase of the **TaHiTI** process is the documentation of results and handover to other processes.

The threat hunting team must process the results from the execution step and document findings. This documentation must cover the most important results of the hunt, and the conclusions drawn based on those results. The documentation may also have recommendations. Recommendations may include improvements to preventative measures (from simple configuration changes to architectural changes), recommendations for logging (additional sources, additional details, etc.), recommendations for security monitoring use cases and process recommendations (improvements in vulnerability or configuration management). Finally, the document should have a 'lessons learned' sections that covers how the hunt has helped the hunters to improve. Lessons learned could also be that the hunters have gained valuable insight into parts of the infrastructure. Such insights may ultimately lead to new hunting activities and make subsequent hunts more efficient.

The final activity is handover to other processes. Potential processes that can receive input from the hunting investigation are security incident response, security monitoring, threat intelligence, vulnerability management and others. These processes are show in figure 5.

[7] http://www.activeresponse.org/wp-content/uploads/2013/07/diamond.pdf

[8] http://apt.threattracking.com

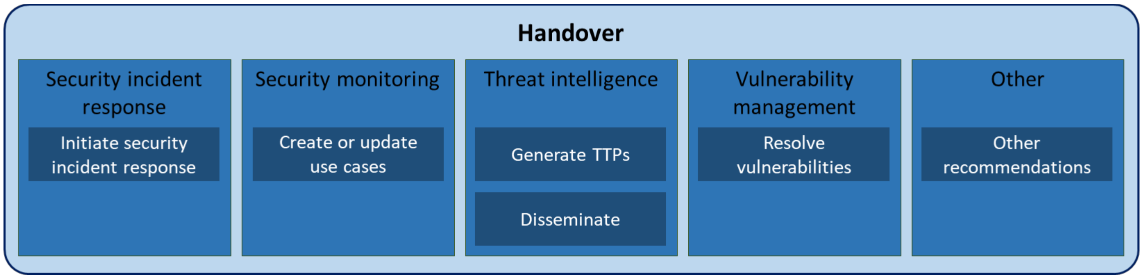

Figure 5: processes triggered by threat hunting investigations

5 Metrics and MaGMa for threat hunting

Metrics are important to determine the efficiency and effectiveness of the threat hunting process and show its added value to the organization. There are 2 basic types of metrics: quantitative (numbers) and qualitative (value). The focus should be on how threat hunting adds value to the organization, so careful selection of metrics is required. The following is a short list of metrics that are indicators of the value added by the threat hunting process:

- **The dwell time of the findings**: since threat hunting should reduce dwell time this should be reported for any compromise uncovered in threat hunting.
- **Incident response**: the number of incidents triggered by the threat hunting process.
- **Security monitoring**: the number of added and updated use cases.
- **Threat intelligence**: new threat intelligence created during the threat hunting process.
- **Security recommendations**: new preventative measures suggested in threat hunting reports.
- **Vulnerability management**: the number of vulnerabilities or misconfigurations uncovered.

When defining metrics for threat hunting, it is important to start out with the goal of the process and then determining useful metrics. While defining metrics that are indications of quality is harder than simply providing numbers, it is well worth the effort.

The **TaHiTI** methodology is supported by the ”MaGMa for threat hunting” tool, which allows hunters to document their results, structure the outcome of their hunting investigations and provide direction for growth of the threat hunting process. Some of the above metrics have been embedded in the MaGMa for threat hunting tool. While this tool can be used separately from the MaGMa Use Case Framework, organizations that are already using MaGMa for their use case management will be able to more easily integrate threat hunting and security monitoring processes. This is due to the common language between these teams and their tools.

6 Conclusion

The **TaHiTI** methodology integrates threat hunting and threat intelligence and provides a clear step-by-step process that hunters can follow to conduct structured hunting investigations. Take these last considerations into account:

1. Carefully select, prioritize and document your input (triggers).
2. Execute hunting with care and apply critical thinking continuously.
3. Use hunting output to drive other security processes and mature and evolve the hunting process itself.

As with any methodology, not all of it may be required for every single hunt. In some cases, hunting investigations will be broad and look at different aspects of a complex TTP. In other cases, hunting investigations may be narrow and scoped at a single specific aspect. Some hunts will benefit from a very formal approach, while others may not. Because each hunt is different, investigations will have different requirements. It is up to the organization to apply the methodology in a flexible way that allows the hunters the freedom to hunt in a standardized and efficient manner, without introducing unnecessary overhead. The threat hunting team should consider which elements are required before initiating a hunt, while retaining the flexibility to apply changes where required.

The full documentation and the MaGMa for threat hunting tool can be obtained from:
https://www.betaalvereniging.nl/en/safety/tahiti/

Authors
Rob van Os, de Volksbank, *lead author*
Marcus Bakker, Rabobank*, co-author*
Ruben Bouman, ING / FinancialCERT
Martijn Docters van Leeuwen, Rabobank
Marco van der Kraan, Rabobank / FinancialCERT
Wesley Mentges, de Volksbank
Armand Piers, ABN AMRO Bank / FinancialCERT

https://www.linkedin.com/company/threathunting
https://www.twitter.com/threathunting_