jamk.fi

Evaluation of Threat Information Feeds for a Cyber Defense Center

A Case Study for Company Netcloud AG

Kuusenmäki Juha

Master’s thesis
October 2020
Technology
Master's Degree Programme in Information Technology
Cyber Security

Jyväskylän ammattikorkeakoulu
JAMK University of Applied Sciences

jamk.fi

Description

Author(s)	Type of publication	Date
Kuusenmäki, Juha	Master's thesis	October 2020
		Language of publication: English
	Number of pages 73	Permission for web publication: x

Title of publication

Evaluation of Threat Information Feeds for a Cyber Defense Center

A Case Study for Company Netcloud AG

Degree programme

Master's Degree Programme in Information Technology, Cyber Security

Supervisor(s)

Saharinen Karo; Hautamäki Jari

Assigned by

Andreas Renold, Netcloud AG

Abstract

Companies and organizations nowadays need to be prepared to defend themselves against more sophisticated, well coordinated and capable cyber-attacks. In order to defend against these attacks more efficiently, companies and organizations need to improve their situational awareness about the emerging threats their businesses are facing. The growing need to improve the situational awareness created a need for Netcloud AG to research threat information feeds suitable for their Cyber Defense Center and their customer and to research their possible benefits for improving situational awareness.

The purpose of this study is to find out what STIX/TAXII threat information feeds currently exist and what are their benefits and disadvantages in addition to find out if the data provided by those feeds can be used to improve situational awareness of an organization.

The study was conducted as a literature review about existing STIX/TAXII feeds which produced an overview of the capabilities, weaknesses and possible use cases of each feed. The results of the literature review were validated with a practical case study. The research results can be used to support Netcloud Cyber Defense Center's further evaluations for threat information feeds suitable for their business strategy and their customer cases.

Keywords/tags (Threat Data, Cyber Threat Information, Cyber Threat Intelligence)

Miscellaneous (Confidential information)

jamk.fi

Kuvaus

Tekijä(t)	Julkaisun laji	Päivämäärä
Kuusenmäki, Juha	Opinnäytetyö, ylempi AMK	Lokakuu 2020
		Julkaisun kieli: Englanti
	Sivumäärä 73	Verkkojulkaisulupa myönnetty: x

Työn nimi

Uhkadatasyötteiden arviointi kyberturvakeskukselle

Tapaustutkimus yritykselle Netcloud AG

Tutkinto-ohjelma

Master's Degree Programme in Information Technology, Cyber Security

Työn ohjaaja(t)

Saharinen, Karo ja Hautamäki Jari

Toimeksiantaja(t)

Andreas Renold, Netcloud AG

Tiivistelmä

Nykypäivänä yritykset ja organisaatiot joutuvat puolustautumaan yhä hienostuneempia, hyvin koordinoituja ja erittäin kyvykkäitä kyber-hyökkäyksiä vastaan. Jotta näitä hyökkäyksiä vastaan voitaisiin puolustautua tehokkaammin, täytyy yritysten ja organisaatioiden parantaa tilannekuvaansa heihin kohdistuvista uusista uhista. Lisääntynyt tarve parantaa tilannekuvaa loi tarpeen tutkia Netcloud AG:n asiakkaille ja kyberpuolustuskeskukselle sopivia uhkadataa tuottavia syötteitä ja niiden mahdollisesti tuomia etuja tilannekuvan parantamiseksi.

Tutkimuksen tarkoituksena oli selvittää mitä uhkadataa tuottavia STIX/TAXII syötteitä on tällä hetkellä olemassa, kartoittaa niiden tuomat edut ja heikkoudet ja selvittää voiko niiden tuottamaan dataa hyödyntää tilannekuvan parantamiseksi. Tutkimus suoritettiin kirjallisuustutkielmalla olemassa olevista STIX/TAXII syötteistä ja todentamalla kirjallisuustutkielman tuloksia käytännössä tapaustutkimuksella.

Kirjallisuustutkimuksen tulosten pohjalta saatiin hyvä yleiskuva olemassa olevista uhkadataa tuottavista STIX/TAXII syötteistä, niiden ominaisuuksista, heikkouksista ja mahdollisista käyttökohteista. Kirjallisuustutkimuksen tulokset todennettiin tapaustutkimuksella. Tutkimustuloksia voidaan hyödyntää Netcloud AG:n kyberturvakeskuksen arvioidessa heidän liiketoimintastrategiaan ja heidän asiakkaille sopivia uhkadataa tarjoavia syötteitä.

Avainsanat (Uhkadata, Kyberuhkatieto)

Muut tiedot (salassa pidettävät liitteet)

Acronyms

AI	Artificial Intelligence
AIS	Automated Indicator Sharing
API	Application Programing Interface
CDC	Cyber Defence Center
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilites and Exposures
C&C	Command and Control
DHS	Department of Homeland Security
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
IDS	Intrusion Detection System
IOC	Indicator of Compromise
IP	Internet Protocol
ISC	Internet Strom Center
JSON	JavaScript Object Notation
MHN	Modern Honey Network
MUTEX	Mutual Exclusion Object
OASIS	Organization for the Advancement of Structured Information Standards
OTX	Open Threat Exchange
PII	Personal Identifiable Information
PKI	Public Key Infrastructure
SA	Situational Awareness
SIEM	Security Information and Event Management
SDO	STIX Domain Object
SRO	STIX Relationship Object
STIX	Structured Threat Information Expression
TAXII	Trusted Automated eXchange of Indicator Information
TIP	Threat Intelligence Platform
TTP	Tactics, Techniques and Procedures
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
XML	Extensible Markup Language

Contents

Figures

Tables

1 Introduction

You might often hear that cyber security is kind of a cat and mouse game between the defender organization and the threat actor. This analogy reminds me of a friend of mine who was struggling with some especially clever mice who kept on invading his warehouse, despite the numerous traditional mousetraps he had set on the floor, leaving only empty traps without a cheese bait behind. He could not pinpoint the holes where the mice came in during the night and he was unsure what they were after. After desperately trying out different ways to dealt with the mouse problem for a few days, his neighbor lent him a trap that worked a bit differently than the traditional traps. This trap had worked well for his neighbor and sure enough, it solved the problem for my fried as well. In this analogy, my friend in the defender role is reactively trying to fend off continuously invading threat actors, who in this case are the mice. What if my friend in the defender role would have proactively received crucial information from his neighbor whose warehouse had been raided by mice just a few days ago? He could have received information how these specific mice had entered their warehouse, what had been chewed on, what kind of traps had been efficient and what had made it possible for the mice to enter their warehouse in the first place. My friend would have been able to make the necessary checks and preparations to prevent the mice from entering his warehouse and chewing on his precious extra old cheese he had been saving for the next fondue evening.

Organizations now days have to defend their sensitive data and systems against increasingly sophisticated cyber-attacks performed by coordinated and capable threat actors. (Johnson, Badger, Waltermire, Snyder, Skorupka 2016). As the threat actors are becoming more capable of causing severe damage to organization's resources, reputation and other important assets it becomes increasingly important to understand the threats and threat actors targeting the organization. An organization also needs to understand the situation of their own assets to be able to make business related or mission related decisions, which can be accomplished with Situational Awareness (Kokkonen, 2016).

An organization can significantly improve its situational awareness and security posture by consuming cyber threat information. The consumed cyber threat information can be processed further into actionable cyber threat intelligence by correlating and enriching the threat information feeds for example by feeding the data into a Security Information and Event Management (SIEM). This makes it possible for an organization to make fast business related decision and to react quickly to relevant emerging threats (Nist, 2016).

2 Research motivation, objective, method and scope

2.1 Research motivation and objective

Situational awareness consists of three factors according to Mitre. These factors are network awareness, mission awareness and threat awareness. Threat awareness is an important factor, which can be improved by identifying internal suspicious behavior, gaining knowledge of external threats and taking part in threat sharing communities (Mitre. 2020). The motivation for this research originated from the Netcloud Cyber Defense Center's and their customers' need for improved threat awareness in their businesses. The objective of this research is to evaluate different existing threat information feeds producing threat data in order to determine which cyber threat information feeds and services could support the current customers and Cyber Defense Center of Netcloud to improve their threat awareness. The evaluation is based on criteria such as event quality, event timeliness, ease of use, event scope and cost. The second objective is to test some of the evaluated feeds in practice to see if the actual data provided by threat information feeds could be beneficial and supports the theoretical evaluation results. Netcloud can use the results of this research to support their evaluation of suitable threat information feeds for their business strategy to provide customers a threat intelligence service for better threat awareness. Different commercial and non-commercial products will be included in the evaluation.

2.2 Research method

The research is based on qualitative research method performed as a literature

review followed by a case study combining both qualitative and quantitative data in a mixed method research. The qualitative research method performed as a literature review was chosen due authors insufficient previous knowledge in the topic and to support the case study by first gaining deeper knowledge in the topic so that the key features and the research questions for the case study could be identified. The case study performed as a mixed method research was chosen to validate the results of the literature review by analysing real-life data.

2.2.1 Qualitative research

Kananen explains that a qualitative research is suitable when the topic of the research is not well known or researched before. In a qualitative research the researcher's goal is to gain good knowledge of the topic and to understand what are the factors in the topic and how they are related (Kananen, 2017, 32-33.)

The qualitative research method for the literature review was chosen to gain a good understanding of existing products producting threat data and to gain knowledge of their differences, shared commonalities and key features.

2.2.2 Literature review

Lauf and Kuziemsky from University of Victoria explain that a literature review is essential to identify existing literature on a subject, determining to what extent a specific research area identifies trends or patterns, aggregating empirical results to define research questions for evidence-based practice and identifiying questions that require further research in the topic (Lauf, F., Kuziemsky, C. 2017).

The literature review research method in this thesis will be conducted first by analysing published literature such as documentation and existing researches about products producing threat data. The collected information will be evaluated and summarized into easily accessible format to provide knowledge about the existing products for both the author and the assigner of this thesis. The literature review will also help to determine the qualities and reseach questions that need to be researched further in the case study.

2.2.3 Mixed method reseach

Accoding to Dr. Roslyn Cameron from Deakin University a Mixed method research involves a philosophical assumption that acts as a guideline for collecting and analyzing both qualitative and quantitative data in a single study or in multiple studies (Cameron, 2015).

A Mixed method research was chosen for the collective case study as it allows mixing both qualitative and quantitative data in a collective case study.

2.2.4 Collective case study

The case study will be conducted as a collective case study combining both qualitative and quantitative data to form a collective understanding of the research questions by studying three different cases (Simons, 2009, 21).

2.3 Research questions

The research questions for this thesis were chosen to provide beneficial data for Netcloud:

- What STIX/TAXII threat information services currently exist?
- What are the benefits and weaknesses of the existing STIX/TAXII services?
- Can consuming of a STIX/TAXII threat information feed help to improve situational awareness in Netcloud's Cyber Defense Center and their customer environments?

2.4 Scope

Scope of the research is limited to cover only a specific set of threat information feeds in STIX format delivered with TAXII protocol. Not all of the existing feeds are covered in this research. The scope of the analyzed fields of the threat information data in the case study is limited to cover only the indicator, indicator type and time stamp fields. Analysis of other fields such as Confidentiality, Severity and TLP are not covered in this research.

2.5 Structure of the Thesis

The thesis starts with an introduction Chapter 1. The introduction Chapter 1 is

followed by Chapter 2, which explains the research background. Chapter 3 goes through the theory of exchanging cyber threat information concentrating mainly on STIX and TAXII standards. Chapter 4 contains the literature review for the evaluation of the exiting threat information feeds. Chapter 5 contains the case study, which tests the measureable qualities of threat information feeds in practice. Chapter 6 presents and correlates the results of the literature review and the case study. Chapter 7 examines and presents the results of the literature review and the case study in more readable format and answers the research questions. Lastly the Chapter 8 contains the conclusions and further development ideas. The Figure 1 illustrates the different stages of the research.

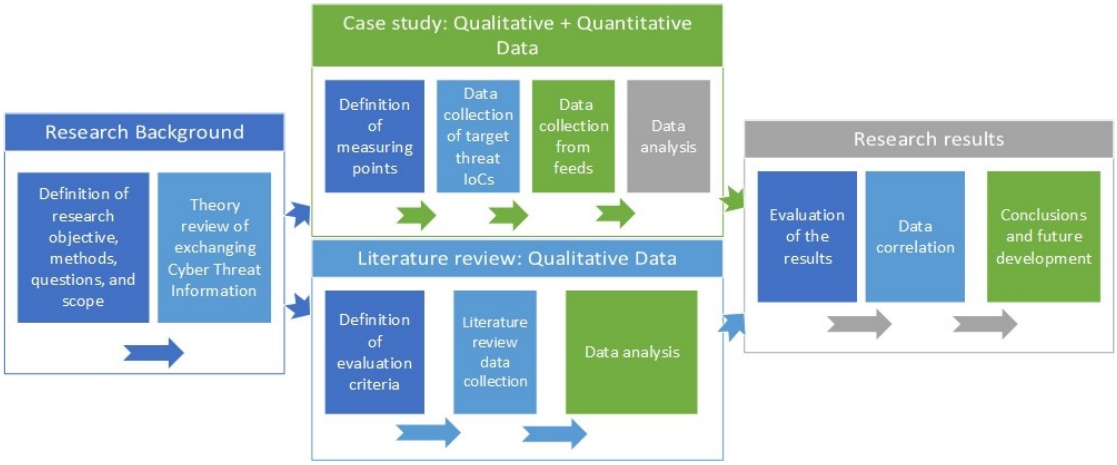

Figure 1. Research structure

3 Exchanging Cyber Threat Information

Any information that can be used to identify and respond to a cyber threat, can be considered cyber threat information (Nist, 2016). Cyber threat information can for example include different types of indicators, observables, incidents, information about a campaign or threat actor, tactics, techniques and procedures (TTPs) (see Figure 2).

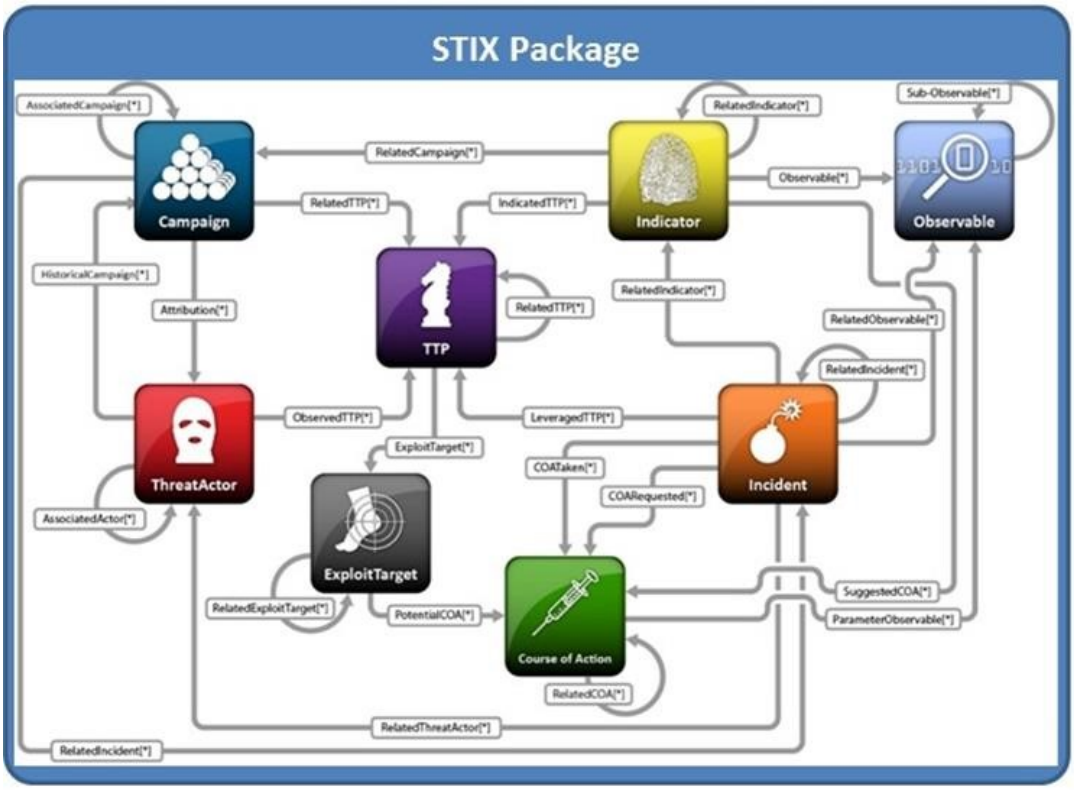

Figure 2. STIX Package (Oasis CTI 2015)

Organizations can benefit from cyber threat information exchange by correlating and analyzing threat information provided by different sources. This cyber threat information can be enriched into more accurate and actionable threat intelligence which can help an organization to understand the threats better to detect possible campaigns targeting their businesses and to react to the threats accordingly (Nist, 2016).

Through gathering threat intelligence an organization can gain rich evidence based intelligence which can be used to support decision making. The accuracy of the threat intelligence relies heavily on the skills, maturity and sources of the entity responsible for the analysis of the threat information (Garrido-Pelaz, González-Manzano, Pastrana 2016).

Indicators can be used to detect ongoing attacks or to investigate if a breach has already happened. Indicators consist of one or multiple different observables such as Internet Protocol (IP) address of a malicious actor, a specific hash of malicious code or an Uniform Resource Locator (URL) containig malicious content (Nist, 2016).

Tactics, techniques and procedures (TTPs) represent the behavior of a cyber actor. TTPs describe what and how a cyber actor does in detail (Nist, 2016). Tactics describe the behaviour in high-level. For example tactics could contain information that a cyber actor uses certain malware to encrypt victims data. Technique could describe at lower-level of details that the malware is delivered by a malicious crafted email which is downloaded and executed when a user opens a malicious link contained in the crafted email. A procedure could for example describe that the victims are chosen by social engineering, the crafted email is sourced from a newly registered domain xyz.com and the command and control connection is opened to specific IP using DNS tunneling to avoid detection.

By exchanging cyber threat information an organization can glue together these breadcrumbs of threat information that might not otherwise be available to the organization to form a better understanding of the threat in hand. The more different components of the threat are identified the better the organization can react to the threat (Nist, 2016). One organization might detect the threat but another organization might be able to use that information to mitigate the threat in their environment. The shared situational awareness and improved security posture are the key benefits of exchanging cyber threat information (Garrido-Pelaz, González-Manzano, Pastrana 2016).

Cyber threat information can be shared over various platforms and protocols in many different formats. Organizations participating in threat intelligence sharing need to form a community of trust to work together for a common goal. Standards for representation and sharing of cyber threat information need to be defined to be able to share cyber threat information (Al-Ibrahim, Mohaisen, Kamhoua, Kwiat, Njilla 2017). A study of Threat Intelligence Sharing Platforms was conducted in 2017 by C. Sauerwein, C. Sillaber, A. Mussman and R. Breu which concluded that the Structured Threat Information Expression (STIX) was the most used standard for sharing threat intelligence in 2017 (Garrido-Pelaz, González-Manzano, Pastrana 2016). This thesis concentrates on cyber threat information exchange in STIX format shared over Trusted Automated eXchange of Indicator Information (TAXII) feeds.

3.1 STIX

STIX is a standardized language and format for representing and exchanging cyber threat intelligence (Oasis, 2017). STIX was developed by the Deparmet of Homeland Security and MITRE. The specification was transitioned to the non-profit organization called Organization for the Advancement of Structured Information Standards (OASIS) to make sure that the STIX/TAXII specifications reamains available to everyone free of charge (Homeland Seurity, 2015).

There are currently two version of STIX available. Version 1.x relies on XML format whereas STIX version 2.x relies on JSON format which makes the STIX 2.x more simplier to understand. STIX 2.0 was significantly re-designed from earlier version. STIX 2.0 adopted only the necessary objects to fulfilll the basic requirements for cyber threat information sharing and the rest of the objects and properties of STIX 1.X were left out (Oasis CTI, 2017).

STIX 2.0 does not integrate objects into other objects like in older version. STIX 1.x for example expresses relationships patterns with XML syntax by integrating the relationship pattern directly into the objects. This makes the STIX 1.x less agile as defining a new relationship between two objects would require editing one of the orignal objects. STIX 2.0 overcomes this restriction by introducing a top-level relationship object (see figure 3.).

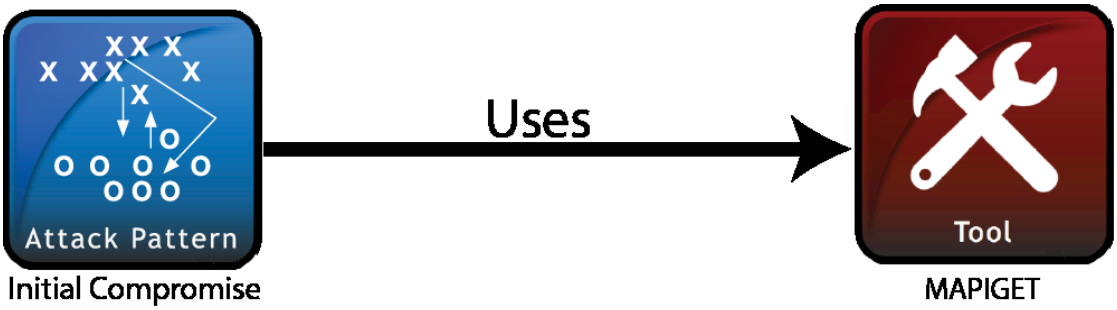

Figure 3. STIX Relationship (Oasis Open)

The top-level relationship objects makes it possible to define a relationship between two objects without having to edit an existing object as shown in the following example of STIX 2.0 relationship object. The relationship object "relationship—xyz" indicates that the attack pattern object "attack-pattern-xyz" uses the the tool "tool-

xyz" to perform the pattern described in the attack pattern object "attack-pattern-xyz".

```
{
  "type": "relationship",
  "id": "relationship--xyz",
  "created": "2020-03-17T22:28:28.1230000",
  "modified": "2020-03-18T21:13:14.1230000",
  "relationship_type": "uses",
  "source_ref": "attack-pattern-xyz",
  "target_ref": "tool-xyz"
}
```

STIX 1.0 relationship is harder to read and less agile as the relationship object is integrated into another object. The following example describes the relationship of the indicator "abc" to the TTP "def". The indicator and the TTP objects are top-level objects but the actual relationship object is integrated into the indicator object "abc".

```
<stix:Indicators>
 <stix:Indicator id="abc" xsi:type='indicator:IndicatorType' timestamp="xyz">
  <indicator:Indicated_TTP>
   <stixCommon:Relationship>Indicates Malware</stixCommon:Relationship>
   <stixCommon:TTP idref="def" />
  </indicator:Indicated_TTP>
 </stix:Indicator>
</stix:Indicators>
<stix:TTPs>
 <stix:TTP id="def" xsi:type='ttp:TTPType' timestamp="xyz"/>
</stix:TTPs>
```

STIX 2.0 specification defines a specific set of STIX Domain Objects (SDOs). STIX 2.0 uses the SDOs as building blocks that can be glued together with other SDOs or with STIX Cyber-observable Objects (SCOs) by utilizing embedded relationships and STIX Relationships Objects (SROs) to create and share in-depth Cyber Threat Intelligence (Oasis CTI, 2017). SDOs specified by STIX 2.0 are:

- Attack Pattern
- Campaign
- Course of Action
- Identity
- Indicator
- Intrusion Set
- Malware
- Observed Data

- Report
- Threat Actor
- Tool
- Vulnerability

STIX 2.1 is the latest STIX version introducing six additional SDOs (Oasis CTI, 2020). Additional SDOs specified by STIX 2.1 are:

- Grouping
- Infrastructure
- Location
- Malware Analysis
- Note
- Opinion

Attack Pattern SDO describes in detail how the threat actors are trying to perform an attack on the target. This object is used to categorize and to map attacks to certain patterns. Attack Pattern SDO describes the pattern in textual format but it can also ference to external non-STIX information such as Common Attack Pattern Enumeration and Classification (CAPEC) Id (Oasis CTI, 2020). The following example of an Attack Pattern SDO describes a Code Injection Attack Pattern ID "xyz" with an external CAPEC reference of 242.


```
{
 "type": "attack-pattern",
 "spec_version": "2.1",
 "id": "xyz",
 "created": "2020-10-16T22:19:00.0001",
 "modified": "2020-10-16T08:17:27.0001",
 "name": "Code Injection",
 "description": "xyz",
 "external_references": [
  {
   "source_name": "capec",
   "external_id": "CAPEC-242"
  }
 ]
}
```


Campaign SDO groups together specific activities performed against specific targets over a certain time period. A Campaign SDO usually describes the objectives and the motivation behind the campaign along with the potential targets and the resources behind the campaign.

Course of action SDO describes an action needed to repond to an ongoing attack reactively or to prevent an attack from happening proactively with specifical technical actions such as patch installation or firewall rule implementation or with higher level actions such as company policy changes. The Course of action SDO describes the action in textual format only in the current STIX version. The action property has been reserved to provide automated courses of actions in the future versions of STIX (Oasis CTI, 2020).

Identity SDO is used to provide information of different identities such as threat actors, attack targets and sources of information ranging from individuals to large organizations or specific groups (Oasis CTI, 2020).

Indicator SDO describes a pattern that can be used to indicate malicious activity. The pattern may be presented as a STIX patterning language following STIX specifications or it can be presented in another language such as Snort, Yara, Sigma, Suricata or Perl Compatible Regualr Expressions language (PCRE). Indicator SDO must contain atleast the type string presenting the indicator, pattern of the indicator presented in appropriate language, pattern type to identify which language is used to present the pattern and the time from which the indicator is considered valid in addition to the common required properties such as spec_version, id, created and modified properties. Indicator SDO can optionally indicate for example to which kill chain phase the Indicator relates or it can provide further information with help of relation ship properties to indentify for example to which campaing or malware the specific indicator relates (Oasis CTI, 2020). Figure 4. Illustrates a malicious activity indicator in STIXX language with an indicator relationship to the X4z9arb Backdoor malware.

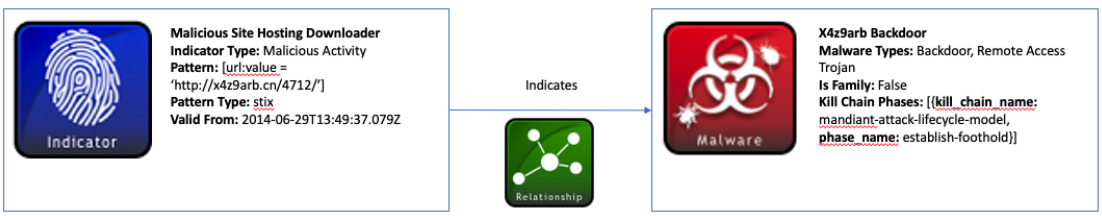

Figure 4. Indicator SDO (Oasis Open)

Intrusion Set SDO groups together multiple malicious activities such as indicators, tools and TTPs with common properties, which can be related to a single threat

actor. Intrusion Set differs from a Campaign SDO in a way that a Campaign groups together specific activities performed against a specific targets over certain time period with a certain motivational driver. The Intrusion Set is a set of different attack methods that may be observed over a long but unspecified time as part of multiple campaigns with multiple motivational drivers (Oasis CTI, 2020).

Malware SDO describes a malware in details by categorizing the malware type and how the malware works by identifiying the malware family, aliases, kill chain phases, operating systems affected by the malware, architecture that allows the execution of the malware, implementation language, capabilities and a list of known identifiers associated with the malware. Malware SDO provides valuable information about malware but it does not provide a pattern presented in a patterning language such as STIX patterning language so that it could be automatically processed by supporting security devices as opposed to Indicator SDO (Oasis CTI, 2020).

Observed Data SDO utilizes data provided by one or multiple STIX Cyber observable Objects to produce intelligence by conveying the raw data of the SCOs such as file hashes or network connections (Oasis CTI, 2020).

Report SDO contains threat intelligence grouped together regarding one or more subjects in a representative manner. Report SDO could be used for example to represent a specific Campaign in a textual format but the details would be included by referencing all the realted SDOs (Oasis CTI, 2020).

Threat Actor SDO represents an entity which is believed to be a part of malicious activity. The entity could be an individual, a group or an organization. The Threat Actor SDO describes the type, known aliases, known roles, motivational goals, skills and resources of the Threat Actor (Oasis CTI, 2020).

Tool SDO represents the properties of legitimate software that could be used in a cyber attack by a Threat Actor. The Tool SDO describes the name, description, type, known aliases, kill chain phases and version of the tool. The information provided by Tool SDO can be used for example to assess the potential usage of the tool during an attack (Oasis CTI, 2020).

Vulnerability SDO is primarily used to provide a link for other SDOs to external non-STIX information about the known vulnerability such as Common Vulnerabilities and Exposures (CVE) or to provide information about 0-day vulnerabilities when a specific vulnerability is being exploited by malicious activity (Oasis CTI, 2020).

Grouping SDO is used to group together STIX Objects that are somehow related by their context. Grouping SDO could be used for example during a cyber attack to quickly group togheter all the known SDOs related to an incident so that this information could be collaborated with other parties affected by the activity. The Grouping SDO can be used to list all the related SDOs, SCOs and SROs to provide a context and a description of the grouped objects (Oasis CTI, 2020).

Infrastructure SDO describes systems, software, services etc. which serve somekind of a role in malicious activity on the attacker, defender or victim side. The Infrastructure SDO contains the name, description, type, known aliases, kill chain phases and time stamps related to an Infrastructure involved in malicious activity (Oasis CTI, 2020).

Location SDO is used to describe the geographical location such as region, country or/and longitude and latitude. The Location SDO is used to provide context to other SDOs (Oasis CTI, 2020). The Location SDO could for example be used to indicate that the Threat Actor APT29 is known to originate from Russia.

Malware Analysis SDO is used to present the results and metadata of a malware analysis performed on a malware. The Malware Analysis SDO presents properties regarding the testing environment along with the systems, software versions, configurations, time stamps and many more optional properties used to analyse the malware. A list of identified SCOs along with the results can be presented in Malware Analysis SDO (Oasis CTI, 2020).

Note SDO is used mostly by human entities to provide context in a form of human readable text. A Note SDO must contain type, content and the STIX object that the note is used for. The Note SDO can optionally contain a summary and the name of the author (Oasis CTI, 2020).

Opinion SDO is used to provide subjective assessment for STIX Objects created by a different entity. A consumer could for example use the Opinion SDO to provide feedback for the producer about an Object that they agree or disagree on. The Opinion SDO could also be used in an organization by an analysts to flag some Indicator SDOs to be considered as a false positive by giving a low enumeration value in the opinion property so that the indicators are not being reacted upon. The required properties of the Opinion SDO are the enumeration value of the opinion property and the STIX object it is referring to. The author of the Opinion SDO can optionally provide a textual explanation and their name. The enumeration value of the Opinion are as follows:

- Strongly-disagree is equivalent to a one out of five
- Disagree is equivalent to a two out of five
- Neutral is equivalent to a three out of five
- Agree is equivalent to a four out of five
- Strongly-agree is equivalent to a five out of five

Clear guidelines for the usage of Opinion SDOs within sharing communities are important as the opinions are subjective and the STIX specification does not provide any best practices for their usage (Oasis CTI, 2020).

Although STIX 2.0 has done some great improvements to STIX 1.x and the STIX 2.1 has improved the language even further by defining new objects and concepts which were covered in the chapter 2.1, there is still room for new improvements. A very usesful feature is the reserved Action property of the Course of Action SDO which will be included in the future versions of STIX. The Action property will make it possibe for machines to automatically react upon the Course of Action SDO which in the current version is not yet possible to acchieve without external automation scripts.

3.2 TAXII

Whereas STIX defines the language and format for cyber threat information representation and exchange, Trusted Automated eXchange of Indicator Information (TAXII) is a protocol operating at the application layer, designed for sharing cyber threat information over hypertext transfer protocol secure (HTTPS). TAXII itself is not an application. It defines a RESTful API with necessary messages and concepts for

sharing cyber threat information. TAXII can be used for sharing data in various formats, but it has been especially designed to be used with STIX format (Oasis CTI. 2018).

TAXII defines two main services to support commonly used information sharing models. These services are Collections and Channels services. Collections make it possible for a producer to host cyber threat information data that can be served to the consumers by request/response method. TAXII clients can use use the Collections interface of an TAXII server to either send data to the TAXII server or to request data from the server (Oasis CTI, 2018). Figure 5. illustrates how a TAXII Collection interfaces operates with consumers.

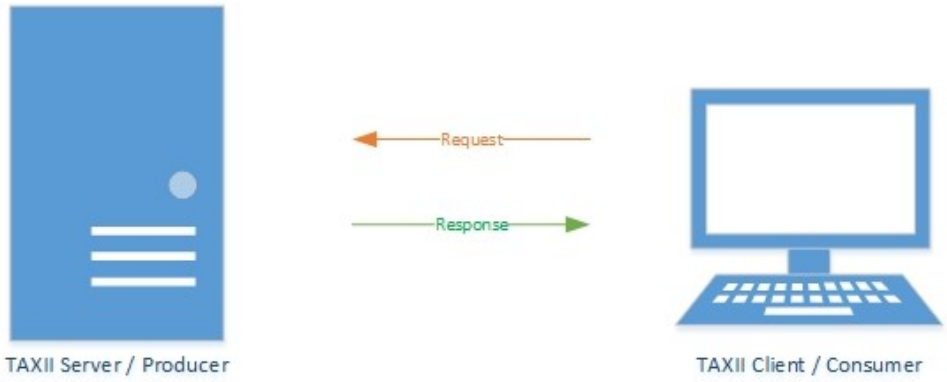

Figure 5. TAXII Collections

TAXII Channels enable producers to publish cyber threat information data to multiple consumers. Consumers can also obtain data from multiple producers. Consumers need to subscribe to a TAXII server that is maintaining a TAXII Channel to be able to receive the data published by the producers. Figure 6. illustrates how TAXII Channels exchange data (Oasis CTI, 2018).

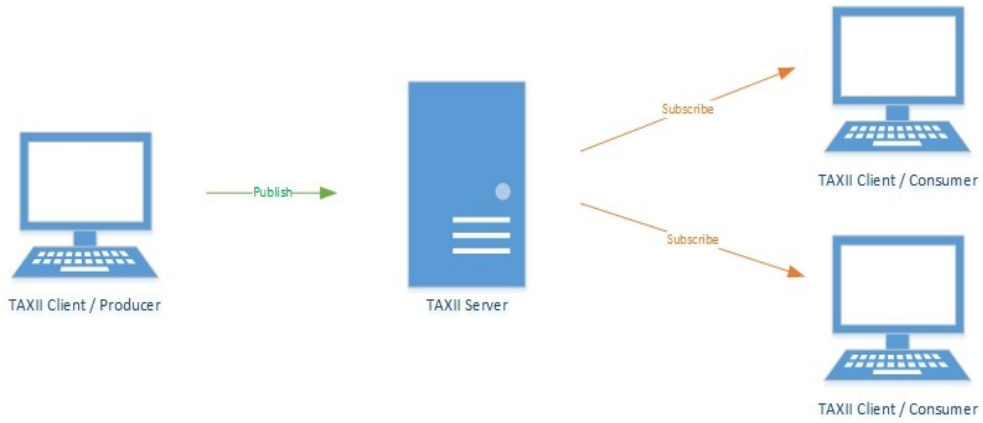

Figure 6. TAXII Channels

TAXII supports commonly used information sharing models hub and spoke, peer to peer, source-subscriber and different variations of all of these information sharing models. In Hub and spoke information sharing model the spokes can share information with the hub, consume information from the hub or both at the same time. In hub and spoke sharing model the data can be sent from hub to spoke and vice versa, but the hub may choose to filter the data received from a spoke before sending the data to spokes (Oasis CTI, 2016b). Figure 7. describes the data flow in hub and spoke architecture.

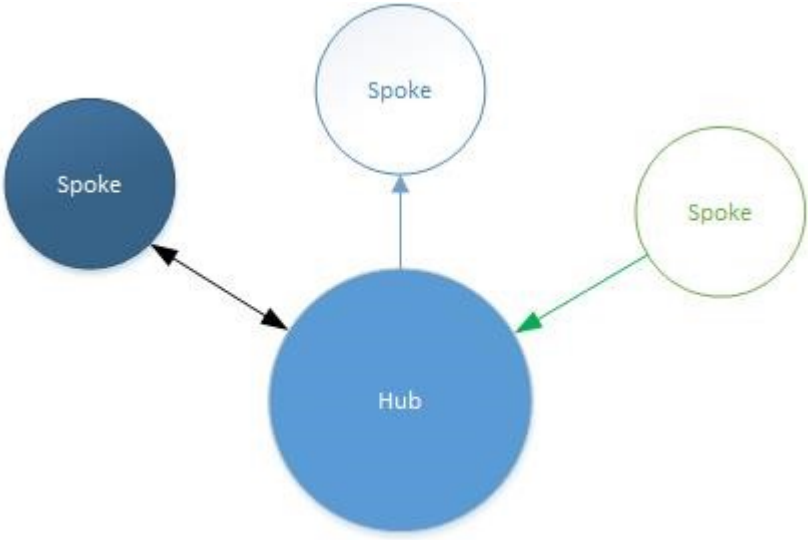

Figure 7. Hub and spoke sharing model

In a Source-subscriber information sharing model the source organization is the only party producing and publishing information for subscribers. In this architecture data can only be sent from the Source to the subscribers. Data can't flow from the subscribers back to the source as shown in the Figure 8. (Oasis CTI, 2016b).

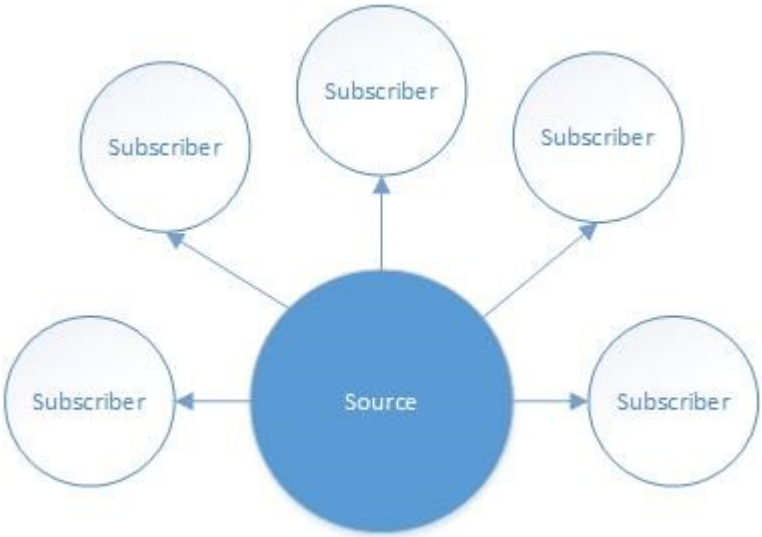

Figure 8. Source-subscriber sharing model

Figure 9. shows the data flow in the peer to peer information sharing model. All the peers can assume the role of a producer and consumer. Data can be sent and received by any peer.

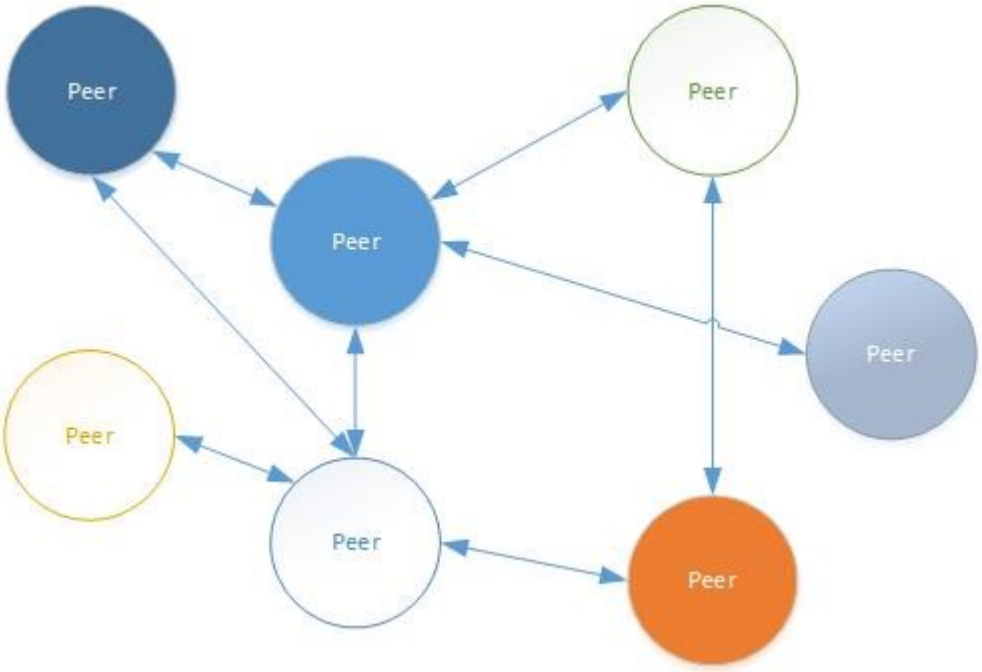

Figure 9. Peer to peer sharing model

4 Literature review

This part of this thesis contains a literature review of Anomali Limo - Emerging Threats Compromised, Anomali Limo - Emerging Threats C&C, Anomali Limo - Dshield Scanning IPs, Anomali Limo - PhishTank, AlienVault OTX, DHS AIS and IBM X-Force Exchange STIX/TAXII threat information feeds. The above mentioned threat information feeds were chosen for this research to gain knowledge of both free to use and commercial feeds provided by both private and public sectors. The review was conducted by analysing published literature of each feed. The feeds were evaluated based on their event quality, event timeliness, ease of use, event scope and cost. The evaluation was summarized into easily accessible format which is introduced in the Table 11 in the Chapter 6.1.8.

4.1 Evaluation criteria and scoring

Common evaluation criteria and scoring values needed to be defined to be able to evaluate the different feeds as objectively as possible relying mostly on documented

numerical values. Some features such as Ease of use and Event quality could not be evaluated entirely objectively based on documented numerical values and the results are therefore subjective opinions of the author of the thesis. A scoring system of three tiers (Excellent, Good and Poor) was applied for all the evaluated criteria to keep the evaluation simple. If some criterion could not be scored based on the literature review, a value of N/A was assigned to that particular criterion. A scoring system with three tiers was not aiming to be as accurate as an atomic clock. The goal of the scoring system was to provide scores for each feed to be able to compare their characteristics to provide a high level overview in easily accessible format and to be able to choose the suitable feeds for more detailed analysis in the case study.

4.1.1 Event Quality

Event quality evaluates the creditability of the events published in the feed. The higher the chance for false positives the lower the event quality of the feed is. The quality is evaluated by examining the sources the threat data is coming from and if the threat data is filtered for false positives by a human entity.

Excellent

A feed would score excellent in Event quality evaluation if the threat data is produced by trusted sources and filtering is done by a human entity to minimize false positives. This type of threat data is almost free of false positives and it can be used as a source for automated incident response.

Good

A feed would score good in Event quality evaluation if the threat data is produced by trusted sources but no filtering for false positives is done or the threat data is produced by non-trusted sources but filtering is done by human entity to minimize false positives. This type of threat data could potentially contain false positives but the rate would be low. The threat data could be used as source for incident response with human supervision.

Poor

A feed would score poor in event quality evaluation if the threat data is produced by non-trusted sources and no filtering is done by human entities to minimize false positives. The rate of false positives can be high in this type of threat data. The data can be used as supporting data to correlate and enrich data extracted from other sources.

4.1.2 Event timeliness

Event timeliness evaluates the time interval how often the new events are published into the feed. The faster the interval is the better the response time to new emerging threats will be.

Excellent

A feed would score excellent in Event timeliness evaluation if new events are published into the feed in real-time or near real-time with maximum delay of two hours. Excellent timeliness of a feed makes a quick reaction time against new emerging threats possible.

Good

A feed would score good in the Event timeliness evaluation if new events are published into the feed with a delay of 2 to 24 hours. This type of threat data is not as current as real-time data, but it is still useful for incident management.

Poor

A feed would score poor in the Event timeliness evaluation if new events are published into the feed with a delay longer than 24 hours. This type threat data can be somewhat useful for incident management, but relying solely on this type of threat data in incident management would make the response times slow. This type of threat data could be helpful for example for forensics analysis.

4.1.3 Ease of use

Ease of use evaluates how easy it is to access the feed and to consume the data provided by the feed. An easy to use feed does not require much expertise to

onboard and to process the threat data further but a feed with complex registration and format might require experts of multiple different areas making it less attractive or even unusable for some organizations.

Excellent

A feed would score excellent in Ease of use evaluation if the feed can be consumed without the need of complex registration procedure and it supports the latest versions 2.0 or 2.1 of STIX. This type of feed is easy to onboard and the registration procedure requires no longer than 24 hours to be completed. The threat data in STIX 2.x formats provide many benefits to older STIX version 1.x.

Good

A feed would score good in Ease of use evaluation if the feed needs registration procedure of medium complexity before it can be consumed and it supports at least version 2.0 of STIX. This type of feed might require some work to onboard but the registration procedure requires no longer than 3 days to be completed. The threat data in STIX 2.0 format provides many benefits to older STIX version 1.x.

Poor

A feed would score poor in Ease of use evaluation if the feed needs a complex registration before it can be consumed or it supports only version 1.x of STIX. This type of feed might require excessive work to onboard and the registration procedure requires more than 3 days to be completed. The threat data in STIX 1.x format lacks the benefits of newer version of STIX.

4.1.4 Event Scope

Event scope evaluates the coverage of the different types indicators of compromise (IoCs) provided by the feed. A broad scope of different types of IoCs is considered beneficial over a narrow scope in this evaluation as being able to detect multiple types of IoCs makes it possible to react to different types of events thus resulting in a better situational awareness.

Excellent

A feed would score excellent in Event scope evaluation if the feed contains indicators of compromise of ten or more different types. This type of threat data would provide a broad set of indicators and the data is highly usable for incident management and response for different types of incidents.

Good

A feed would score good in Event scope evaluation if the feed contains indicators of compromise of at least five different types. This type of threat data would still be usable for incident management and response but the scope would be narrow.

Poor

A feed would score poor in Event scope evaluation if the feed contains indicators of compromise of less than five different types. The scope of this type of threat data is very narrow. It could be useful for automated tasks such as creating firewall rules for known C&C server IPs etc. and it could provide valuable information for example for forensics analysis but as a sole threat data source to provide better situational awareness it is not sufficient.

4.1.5 Cost

Cost evaluates the need for investments required for the setup and usage of the feed. Cost factors of additional human, hardware and software resources are taken into consideration.

Excellent

A feed would score excellent in Cost evaluation if the registration for the service is free of charge and the setup and usage does not require additional human, hardware or software resources.

Good

A feed would score good in Cost evaluation if the registration is subjected to a charge but the setup and usage does not require additional human, hardware or software resources.

Poor

A feed would score poor in Cost evaluation if the registration is subjected to a charge and the setup and usage require additional human, hardware or software resources.

5 Case Study

The purpose of this case study is to gather threat data of four different STIX/TAXII feeds for a test period of two weeks and to compare the threat data with documented indicators of three high profile cyber security threats that were active during the test period to be able to answer the research question: "Can consuming of a STIX/TAXII threat information feed help to improve situational awareness in Netcloud's Cyber Defense Center and their customer environments?"

The case study is carried out by first collecting known indicators for three high profile cyber security threats, which are called target threats in this case study. The known indicators of the target threats are extracted from reports provided by known companies or organizations during the test period between 13.7.2020 and 26.7.2020.

The actual threat data for this case study is collected by consuming the following STIX/TAXII feeds. Other feeds were left out of this case study due to time constraints.

- Emerging Threats Compromised (Limo)
- Emerging Threats C&C (Limo)
- Dshield Scanning IPs (Limo)
- AlienVault OTX

Emerging Threats Compromised, Emerging Threats C&C and Dshield Scanning IPs feeds are all provided by the same Anomali Limo repository. The IoCs of these feeds are however analyzed individually to be able to correlate the results later with the literature review results. AlienVault OTX feed is analyzed as a single feed as it is. Threat data of each of these feeds is consumed between 13.7.2020 and 26.7.2020. The data is exported in CSV-files from STAXX clients for further analysis and comparison.

5.1 Measuring points

The measuring points for the collected threat data needed to be defined to be able measure the data in a way that the results of the case study could be correlated with the results of the literature review and to answer the research question mentioned in the Chapter 5.

Event quality in the literature review was evaluated by examining the credibility of the sources the data is coming from. In the case study it is not possible to measure the creditability of the of the data sources. The case study compares the data of a threat information feed with the documented indicators of each target threat to measure how many of the documented target threat indicators were provided by the threat information feed during the test period. The results will be expressed as numerical and percentual values. The results should provide information if the data provided by the feed would have been useful for an organization when investigating and mitigating an incident related to a target threat. This measuring point will be called Matched indicators in the case study.

Event timeliness will be measured by comparing the publication time stamps of the documented target threat indicators and the first seen time stamps of the matching indicators in the threat information feeds to measure the time gap between the publication time stamp and the first seen time stamp. Same evaluation criteria will be used in the case study for the Event timeliness as defined in the chapter 4.1.2 in the literature review.

Event scope will be measured by examining the total number of indicator types introduced in the threat information feed during the test period of two weeks. Same evaluation criteria will be used in the case study for the Event scope as defined in the chapter 4.1.4 in the literature review.

Ease of use will be measured by examining the supported STIX/TAXII formats and the complexity of the onboarding process. Same evaluation criteria will be used in the case study for the Ease of use as defined in the chapter 4.1.3 in the literature review.

Cost will be measured by examining the financial investments that are required for the onboarding and the usage of the feed. Same evaluation criteria will be used in the case study for the Cost as defined in the chapter 4.1.5

5.2 Test environment

Test environment for the case study was built on two CentOS based virtual machines running on VMware Workstation Pro hypervisor. Two separate instances of STAXX STIX/TAXII client were used to consume the STIX/TAXII feeds and to export the threat data for further analysis. One instance of STAXX was dedicated to AlienVault OTX Feed and the second instance of STAXX was shared by the Emerging Threats Compromised, Emerging Threats C&C and Dshield Scanning IPs Limo feeds.

Hypervisor:

- VMware Workstation 12 Pro
- VMware Workstation 12.1.1 build-3.770944

VM:

- CentOS Linux 7
- 4GB Ram
- *2 Processors*
- *100 GB Hard Disk*

STAXX STIX/TAXII Client:

- *Anomali STAXX Version 3.4.0*
- *Anomali STAXX Build 566*

5.3 Target Security Threats

Three target security threats of high importance were chosen for the case study. The return of the infamous Emotet malware was chosen as the first target threat for the case study due to its high media visibility and notorious reputation over the past few years. The adversary APT29's campaign targeting COVID-19 vaccine development was chosen as the second target threat due to its political nature. The WastedLocker ransomware was chosen as the third target threat due to its noticeable presence in a major Cyberattack against Garmin, which interrupted their online services and encrypted some internal systems.

5.3.1 Target Threat 1. Emotet Comeback

After five months of inactivity the Emotet malware has returned on the radar with new Malspam campaigns. Emotet's activities slowed down and stopped in February 2020 only to re-emerge in July 2020. Checkpoint researchers believe that the developers behind the Emotet botnet were upgrading its features during the down time. Emotet has been seen spreading malspam to infect its targets with Qbot and Trickbot Trojans in July 2020. Emotet has potentially a very high global impact as it is estimated to impact 5% of all organizations worldwide (Checkpoint, 2020).

Emotet is an infostealer malware which has been designed to steal data such as emails, contacts, passwords, transactional data etc. Emotet is usually loaded to the victims computer by a macro hidden in a malicious pdf- , Office- or encypted zip-file attached to an email. The stolen information is uploaded to a command and control server. Emotet uses email for lateral movement by crafting an email to an existing email thread based on the stolen data, thus making the email very convincing. A group called Cryptolaemus is actively tracking the Emotet to provide actual daily indicators related to Emotet (Traficom, 2020).

Emotet operates on three different botnets which are referred as Epoch 1, 2 and 3. Cryptolaemus provided total of 26988 identified indicators related to all three Epochs. The indicators contain known document downloader links, file hashes, payload URLs and C2 IP addresses used by all thee Emotet Epochs. Table 1. displays the volumes of the indicators provided by Cryptolaemus between 13.7.2020 and 26.7.2020 (Cryptolaemus, 2020). All the indicators are included as Appendix 2.

Table 1. Emotet IoCs extracted from Cryptolaemus

IoC type	Relates to	Known IoCs
Mal_ip / C2_ip	Epoch 1 C2s	424
Mal_ip / C2_ip	Epoch 1 Spam C2s	15
Mal_ip / C2_ip	Epoch 1 Stealer C2s	12
Mal_url	Epoch 1 Payload URL	105
Mal_url	Epoch 1 Document Downloader Link	609
Mal_sha256	Epoch 1 Document Payload SHA256	1305

Mal_sha256	Epoch 1 Loader EXE SHA256	7777
Mal_ip / C2_ip	Epoch 2 C2s	425
Mal_ip / C2_ip	Epoch 2 Spam C2s	11
Mal_ip / C2_ip	Epoch 2 Stealer C2s	12
Mal_url	Epoch 2 Payload URL	120
Mal_url	Epoch 2 Document Downloader Link	858
Mal_sha256	Epoch 2 Document Payload SHA256	1457
Mal_sha256	Epoch 2 Loader EXE SHA256	6856
Mal_ip / C2_ip	Epoch 3 C2s	337
Mal_ip / C2_ip	Epoch 3 Spam C2s	6
Mal_ip / C2_ip	Epoch 3 Stealer C2s	12
Mal_url	Epoch 3 Payload URL	85
Mal_url	Epoch 3 Document Downloader Link	327
Mal_sha256	Epoch 3 Document Payload SHA256	1079
Mal_sha256	Epoch 3 Loader EXE SHA256	5156

5.3.2 Target Threat 2. APT29 targets COVID-19 vaccine development

The United Kingdom's National Cyber Security Centre (NCSC) published a joint alert together with Canada's Communications Security Establishment (CSE) and The United States' National Security Agency (NSA) about a new campaign of a Russian cyber espionage group called APT29 targeting various organizations involved in COVID-19 vaccine research on July 16th 2020. The publications states that the adversary APT29, also known as "the Dukes" or "Cozy Bear" is a cyber espionage group, most certainly part of the Russian intelligence service. The motivation behind the campaign might possibly be the intention of stealing information about COVID-19 vaccine research and testing (NCSC, 2020).

APT29 has been performing vulnerability scanning against the public IP-addresses of their target organizations. The group is known to exploit public vulnerabilities such as

- CVE-2019-19781
- CVE-2019-11510
- CVE-2018-13379
- CVE-2019-9670

The group is likely to exploit many more known vulnerabilities once published. Once the group gains access to the target organization, they will most likely try to drop further tools such as WellMess or WellMail and / or to steal legitimate credentials to gain persistent foothold and to send the stolen information to C2 servers (NCSC, 2020).

The joint alert report from NCSC, CSE and NSA provides a total of 93 indicators related to APT29 Activity against COVID-19 vaccine development which contains file hashes and malicious IP addresses as displayed in the Table 2. (NCSC, 2020). All the indicators are included as Appendix 2.

Table 2. APT29 COVID-19 IoCs extracted from the joint alert of NCSC, CSE and NSA

IoC type	Relates to	Known IoCs
Mal_md5	WellMess	32
Mal_ip	WellMess	50
Mal_md5	WellMail	2
Mal_ip	WellMail	1
Mal_md5	SoreFang	3
Mal_ip	SoreFang	1
Mal_ip	Associated with GlobalSign Certificate used by APT29 (WellMail)	4

5.3.3 Target Threat 3. WastedLocker

According to Jim Walter from SentinelOne the WastedLocker ransomware is somewhat new threat which has been on the radar since April 2020 targeting high-value targets in multiple industries. The name WastedLocker originates from the

string "Wasted" which has been seen appended to the encrypted files targeted by WastedLocker. WastedLocker payload can be delivered by multiple different ways. Cobalt Strike and SocGholish frameworks have been seen together with WastedLocker activities for payload delivery and lateral movement. Prevention is as highly important with Wastedlocker as it is with other ransomware to stop the attacker on time (Walter, 2020). Kaspersky explains that the officially confirmed cyber-attack against Garmin in July 2020 causing an interruption of online services and encryption of some Garmin's internal systems was targeted by the WastedLocker (Kaspersky, 2020). Walter has identified eleven SHA256 and eleven SHA1 hashes related to WastedLocker as shown in the Table 3. All the indicators are included as Appendix 2.

Table 3. WastedLocker IoCs extracted from SentinelOne

IoC type	Relates to	Known IoCs
Mal_sha256	WastedLocker	11
Mal_sha1	WastedLocker	11

6 Research results

6.1 Literature review results

This chapter introduces and summarizes the results of the literature review. Data for the literature review was collected for seven different STIX/TAXII threat information feeds by analysing published literature of each feed.

6.1.1 Emerging Threats Compromised (Limo)

Emerging Threats Compromised feed from Anomali's set of threat information feeds called Limo provides information of known compromised hosts, bots and phishing sites. The hosts in this feed are not just regular spam hosts. A host needs to be severely infected or hostile go get listed in this feed. The data is sourced from several highly reliable private data sources (Emerging Threats, 2017a).

Event Quality

Emerging Threats Compromised feed used data produced by highly reliable data sources such as BruteForceBlocker (Emerging Threats, 2017a). The data is not verified or filtered for possible false positives. Due to lack of any filtering or verification of the data, even though sourced from highly reliable sources, the feed could potentially contain a small amount of false positives. Data sourced from highly reliable sources without any filtering results in in a good grade.

Event timeliness

The Emerging Threats Compromised feed is updated on a daily basis (Emerging Threats, 2018). The good timeliness of this feed makes a quick reaction time against new emerging compromised hosts possible.

Ease of use

The Limo service is fully STIX/TAXII 2.0 compliant, which makes the setup very easy by using a TAXII client. The service is well documented, the Anomali Community is very lively and the Anomali provides support for the service over multiple media. The easy setup procedure and the support for STIX/TAXII 2.0 results in an excellent grade.

Event Scope

The Event Scope of the Emerging Threats Compromised feed is very narrow as it provides indicators only for IP addresses of known severely compromised hosts. The data provided by the Emerging Threats Compromised feed could be valuable when used together with data from other sources to correlate and enrich the data. The narrow scope results in a poor result when evaluating the Event Scope of this single feed alone.

Cost

The setup and usage of the Limo feeds are simple and free of charge. The implementation and usage of the feeds do not require extensive additional human, hardware or software resources. Only a STIX / TAXII compliant client is required to consume the feeds resulting in an excellent grade.

Summary

Table 4. Emerging Threats Compromised (Limo) Summary

Evaluation Criteria	Evaluation	Details
Event Quality	Good	Data is sourced from highly reliable sources but no filtering or verification is done.
Event Timeliness	Good	Feed is updated on a daily basis.
Ease of use	Excellent	Fully STIX / TAXII 2.0 compliant and no registration is needed. Service onboarding is well documented and support is provided by Anomali.
Event Scope	Poor	Only IoCs of compromised host IP addresses are provided.
Cost	Excellent	The setup and usage is very easy and free of charge.

6.1.2 Emerging Threats C&C (Limo)

Emerging Threats C&C feed provides information of known active Botnets and other C&C hosts. The data is sourced from nonprofit security organizations Abuse.ch and The Shadowserver Foundation (Emerging Threats, 2018).

Event Quality

The Feodo Tracker operated by a non-profit organization Abuse.ch is considered a highly reliable source. The C&C servers associated with Feodo or some of its evolved variants are tracked and identified by Abuse.ch (Abuse.ch, 2020). The Shadowserver Foundation is another important source of data for the Emerging Threats C&C feed. They use IPv4 scans, sinkholes, large sensor networks consisting of honeypots and honeyclients to collect threat data. The raw data is analyzed by The Shadowserver Foundation (Shadowserver, 2020). The event Quality of the Emerging Threats C&C feed scores excellent as the data is sourced from highly reliable sources and filtering for the data is done by a human entity.

Event timeliness

The Emerging Threats C&C feed is autogenerated and updated on a daily basis (Emerging Threats, 2018). The good timeliness of this feed makes a quick reaction time against new emerging botnets and other C&C hosts possible resulting in a good grade.

Ease of use

The Limo service is fully STIX/TAXII 2.0 compliant, which makes the setup very easy by using a TAXII client. The service is well documented, the Anomali Community is very lively and the Anomali provides support for the service over multiple media. Ease of use scores an excellent grade.

Event Scope

The Event Scope of the Emerging Threats C&C feed is very narrow as it provides IoCs only for IP addresses of known active Botnets and other C&C hosts. The data provided by the Emerging Threats C&C feed could be valuable when used together with data from other sources to correlate and enrich the data. The narrow scope results in a poor result when evaluating the Event Scope of this single feed alone.

Cost

An excellent grade for the cost was given due to same factors as stated in the chapter 6.1.1 for the Emerging Threats Compromised feed.

Summary

Table 5. Emerging Threats C&C IPs Summary (Limo)

Evaluation Criteria	Evaluation	Details
Event Quality	Excellent	Data is produced and analyzed by highly reliable sources.
Event Timeliness	Good	Feed is updated on a daily basis.
Ease of use	Excellent	Fully STIX / TAXII 2.0 compliant and no registration is needed. Service onboarding is well documented and support is provided by Anomali.

Event Scope	Poor	Only IoCs of known C&C host IP addresses are provided.
Cost	Excellent	The setup and usage is very easy and free of charge.

6.1.3 DShield Scanning IPs (Limo)

DShield Scanning IPs feed from Anomali's Limo set sources its data from DShield service, which is a community driven service providing information of IP addresses that have been identified to perform high volume scanning activities. The service is provided by the Internet Storm Center (ISC) (Anomali, 2017). ISC gathers firewall and intrusion detection system (IDS) logs from volunteer community members. The sensors of ISC cover currently over 500.000 IPs in over 50 countries around the globe (Dshield, 2020).

Event Quality

The DShield Scanning IPs feed uses firewall and IDS log data from volunteer users without any sanitation, which could result in false positive indicators. As the data is sourced from untrusted sources and no filtering is done to the data the event quality of the DShield feed is scored poor. The data is not suitable for automated incident management and response but it can be used as supporting data to correlate and enrich data extracted from other sources.

Event timeliness

The DShield Scanning IPs feed is updated nearly in real-time (Dshield, 2020). The excellent timeliness of the DShield Scanning IPs feed makes a quick reaction time against these threats possible resulting in an excellent grade.

Ease of use

All the STIX / TAXII feeds provided by the Anomali Limo service are fully STIX/TAXII 2.0 compliant and no registration is needed (Anomali, 2020). This makes the setup very easy by using a TAXII client. The JSON format of STIX 2.0 makes the data flexible and simple to read. The service is well documented, the Anomali Community is very

lively and the Anomali provides support for the service over multiple media. Ease of use scores an excellent grade.

Event Scope

The Event Scope of the DShield Scanning IPs feed is very narrow as it provides IoCs specifically only for Scanning IPs (Anomali, 2017). The data provided by the DShield Scanning IPs feed is sourced from a very large network of sensors from around the globe covering a large geographical area thus providing a good overview of current scanning activities in the internet but he narrow scope results in a poor result when evaluating the Event Scope of this single feed alone.

Cost

An excellent grade for the cost was given due to same factors as stated in the chapter 6.1.1 for the Emerging Threats Compromised feed.

Summary

Table 6. Dshield Scanning IPs Summary (Limo)

Evaluation Criteria	Evaluation	Details
Event Quality	Poor	Data is produced from untrusted sources without filtering.
Event Timeliness	Excellent	Feed is updated in real-time.
Ease of use	Excellent	Fully STIX / TAXII 2.0 compliant and no registration is needed. Service onboarding is well documented and support is provided by Anomali.
Event Scope	Poor	Only IoCs of Scanning IP addresses are provided.
Cost	Excellent	The setup and usage is very easy and free of charge.

6.1.4 PhishTank (Limo)

PhishTank feed from Anomali's set of threat information feeds called Limo provides information about phishing sites. The PhishTank is operated by OpenDNS but the

phishing data is community based data that can be submitted and verified by the community members (Phishtank, 2020a).

Event Quality

Phishtank feed uses data submitted and verified by multiple community members to avoid false positive indicators. A new Phish will be published after it has been verified by multiple registered community members. The number of entities required for a successful verification depends on the history of the individual voters. A Phish can never be verified just by a single vote (Phishtank, 2020a). The good event quality of the Phishtank threat data is achieved by the community based verification as it minimizes the chance for false positive identifications greatly by performing the necessary filtering for the threat data sourced from non-trusted sources. This makes the data produced by Phishtank feed actionable and results in a good grade.

Event timeliness

Anomali Limo documentation does not contain any information about the update interval of the PhishTank feed. The PhishTank API will allow only a few daily lookups without an API key. With an API key the lookups are not limited but the files containing new Phishes are updated every hour (Phishtank, 2020c). The excellent timeliness of this feed makes a quick reaction time against new emerging threats possible.

Ease of use

All the STIX / TAXII feeds provided by the Anomali Limo service are fully STIX/TAXII 2.0 compliant and no registration is needed (Anomali, 2020). This makes the setup very easy by using a TAXII client. The JSON format of STIX 2.0 makes the data flexible and simple to read. The service is well documented, the Anomali Community is very lively and the Anomali provides support for the service over multiple media. Ease of use scores an excellent grade.

Event Scope

The Event Scope of the PhishTank feed is very narrow as it provides IoCs specifically only for Phishing URLs (Phishtank, 2020b). The data provided by the PhishTank feed

is valuable but the narrow scope results in a poor result when evaluating the Event Scope of this single feed alone.

Cost

An excellent grade for the cost was given due to same factors as stated in the chapter 6.1.1 for the Emerging Threats Compromised feed.

Summary

Table 7. PhishTank Summary (Limo)

Evaluation Criteria	Evaluation	Details
Event Quality	Good	Data is submitted and verified by multiple registered community members to minimize the chance for false positives.
Event Timeliness	Excellent	PhishTank files containing new Phishes are updated every hour.
Ease of use	Excellent	Fully STIX / TAXII 2.0 compliant and no registration is needed. Service onboarding is well documented and support is provided by Anomali.
Event Scope	Poor	Only IoCs for Phishing URLs are provided.
Cost	Excellent	The setup and usage is very easy and free of charge.

6.1.5 AlienVault (AT&T Cybersecurity) OTX Evaluation

Open Threat Exchange (OTX) is a community based threat information sharing and analysis platform. OTX community consists of more than 100,000 participants worldwide and it provides over 19 million indicators of compromise daily. Threat data is shared in OTX to the service subscribers as OTX pulses. These OTX pulses include one to multiple indicators of compromise related to a certain threat (AT&T, 2020a). This evaluation focuses only on the OTX TAXII / STIX feed provided by AlienVault.

Event Quality

The threat data of the AlienVault OTX feed is provided mainly by the OTX community. The data is analyzed and sanitized by the Alien Labs team before it is pushed to the service subscribers as an OTX pulse (AT&T, 2019). The data sanitation done by Alien Labs minimizes the possibility for false positives and enriches the threat data, which makes the threat data produced by AlienVault OTX feed highly actionable and results in a good grade.

Event timeliness

The AlienVault OTX feed operates at near real-time. The small delay between the observation and the pulse is caused by the data sanitation process performed by Alien Labs (AT&T, 2019). The excellent timeliness of the AlienVault OTX feed makes a very quick reaction time against new emerging threats possible.

Ease of use

The AlienVault OTX feed is STIX/TAXII 1.x compliant, which makes the setup very easy by using a TAXII client, but the support for only STIX 1.x lacks the benefits of newer versions 2.0 and 2.1 therefor reducing the grade from excellent down to poor. The service is well documented and the OTX Community is active and assists with questions regarding the OTX service.

Event Scope

The AlienVault OTX feed includes a total of eleven different types of IoCs included (AT&T, 2020b):

- IP addresses
- Domains
- Hostnames (subdomains)
- Email
- URL
- URI
- File Hashes: MD5, SHA1, SHA256, PEHASH, IMPHASH
- CIDR Rules
- File Paths
- MUTEX name
- CVE numberCost

This broad set of indicators supported by the AlienVault OTX feed can provide valuable data while forming a comprehensive understanding of the threat in hand. This single feed covers a broad set of IoCs giving the feed an excellent grade for Event Scope.

Cost

The registration, setup and usage of the AlienVault OTX STIX / TAXII feed are simple and free of charge. The implementation and usage of the feed do not require extensive additional human, hardware or software resources. Only a STIX / TAXII compliant client and an OTX API key is required to consume the feed. The OTX API key can be obtained free of charge by registering to the OTX community resulting in an excellent grade.

Summary

Table 8. AlienVault OTX Summary

Evaluation Criteria	Evaluation	Details
Event Quality	Good	The OTX feed data is provided by the OTX community and the data is sanitized by AlienVault Labs security research team.
Event Timeliness	Excellent	The AlienVault OTX feed operates at near real-time.
Ease of use	Poor	The OTX feed is only STIX/TAXII 1.x compliant. A simple registration procedure is needed to retrieve the OTX API key to be able to authenticate and consume the feed.
Event Scope	Excellent	The OTX feed includes 11 different types of IOCs.
Cost	Excellent	The setup and usage is very easy and free of charge.

6.1.6 Homeland Security – Automated Indicator Sharing (AIS) Evaluation

Automated Indicator Sharing (AIS) is free of charge TAXII service provided by the Department of Homeland Security (DHS) for sharing threat data. The AIS uses data from volunteer organizations participating in the indicator of compromise (IoC)

sharing through AIS. DHS aims to share as many IoCs as possible as rapidly as possible through AIS (Homeland Security, 2020).

Event Quality

The AIS feed uses data from volunteer organizations. The Department of Homeland Security does not validate the indicators of compromise received from the contributors before sharing them in the AIS feed because the strategy of AIS values speed over quality. The AIS service automatically analyzes and removes only personal identifiable information (PII) from the indicator if not directly related to the threat (Homeland Security, 2020). The threat data provided by AIS can be used as additional information to support threat analysis but the community nature and lack of sanitation of the data could potentially produce a high volume of false positive indicators resulting in a poor Event Quality.

Event timeliness

The AIS feed shares the IoCs in real time as soon as a company or federal agency observes an attempted compromise (Homeland Seurity, 2020). The excellent timeliness of the AIS feed makes a very quick reaction time against new emerging threats possible.

Ease of use

The onboarding process of AIS is very complex. AIS requires a PKI certificate from a commercial provider. A static IP-Address needs to be communicated with the Department of Homeland Security and an Interconnection Security Agreement must be signed before being able to consume the AIS feed (Homeland Security, 2020). The very complex onboarding process and the lack of support for STIX 2.0 and STIX 2.1 results in a poor grade for Ease of use.

Event Scope

The AIS feed uses all the indicator types supported by STIX 1.x which results in a very broad Event Scope resulting in an excellent score.

Cost

The setup requires a static IP-Address and a PKI Certificate purchase from a commercial provider which generates costs. The usage is of AIS is free of charge resulting in a good grade.

Summary

Table 9. Homeland Security AIS Summary

Evaluation Criteria	Evaluation	Details
Event Quality	Poor	The AIS feed uses data from volunteer organizations. The Department of Homeland Security does not validate the indicators of compromise shared in the AIS feed.
Event Timeliness	Excellent	The AIS feed shares the indicators of compromise in real time as soon as a company or federal agency observes an attempted compromise.
Ease of use	Poor	The onboarding process of AIS is complex. AIS requires a PKI certificate from a commercial provider. A static IP-Address needs to be communicated with the Department of Homeland Security and an Interconnection Security Agreement must be signed before being able to consume the AIS feed. Only supports STIX 1.x.
Event Scope	Excellent	The AIS feed uses all the indicator types supported by STIX 1.x.
Cost	Good	The setup requires a static IP-Address and a PKI Certificate purchase from a commercial provider which generates costs. The usage is of AIS is free of charge.

6.1.7 IBM X-Force Exchange Evaluation

IBM X-Force Exchange is a commercial cloud-based platform for threat information sharing. IBM X-Force Exchange platform contains different cyber threat intelligence services (IBM, 2020a). This evaluation focuses only on the advanced threat

protection feed of the IBM X-Force Exchange platform. The Advanced threat protection feed provides curated threat information from various sources.

Event Quality

Advanced Threat Protection Feed includes actionable indicators curated by IBM X-Force IRIS team. The feed classifies indicators as actionable if the indicator belongs to a specific threat category such as Scanning IP, Phishing doman, Malware source or some other category and the actionability score reaches a specific high water mark. According to IBM the feed has been tested by external parties to have a detection rate of 99.97% and a false positive rate of 0.003% (IBM, 2020c). This high detection rate and low false positive rate makes the threat data provided by Advanced threat protection feed highly actionable resulting in an excellent grade.

Event timeliness

Information about Event timeliness was not documented by IBM or other sources. This aspect of the Advanced Threat Protection cannot be evaluated.

Ease of use

Onboarding of Advanced Threat Protection Feed requires an API key, which can be obtained by registering for an IBM ID. The registration process is straight forward. The feed is fully STIX/TAXII 1.0 and 2.0 compliant, which makes the setup very easy by using a TAXII client (IBM, 2020b). The service is well documented and IBM provides premium support for this commercial product. The support for STIX 2.0 and the straightforward registration process results in an excellent grade for Ease of use.

Event Scope

Advanced Threat Protection Feed introduces 23 documented different indicator types in 10 different categories (IBM, 2020b):

- Anonymization Services - IPv4, IPv6 and URL
- Botnet CnC Servers - IPv4, IPv6 and URL
- Bots - IPv4 and IPv6
- Cryptocurrency mining - IPv4, IPv6 and URL
- Early Warning – URL
- IRIS – IPv4, IPv6 and URL
- Malware - IPv4, IPv6 and URL
- Phishing - URL
- Scanning IPs - IPv4 and IPv6

- Top Activity - URL / 10K

The very broad Event Scope of Advanced Threat Protection Feed results in an excellent score.

Cost

IBM X-Force Exchange is a commercial product and the usage costs after the test period of 30 days. The cost details were not documented by IBM or other sources. This aspect of the Advanced Threat Protection cannot be evaluated.

Summary

Table 10. IBM X-Force Exchange Summary

Evaluation Criteria	Evaluation	Details
Event Quality	Excellent	Data is curated by IBM X-Force IRIS team and false positive rate of 0.003% has been tested by external parties.
Event Timeliness	N/A	Not available.
Ease of use	Excellent	Registration for IBM ID to obtain an API key is straight forward. The feed is STIX/TAXII 1.0 and 2.0 compliant.
Event Scope	Excellent	The feed includes 23 different types of IOCs in 10 different categories.
Cost	N/A	Not available.

6.1.8 Summary of literature review

The literature review provided an easily accessible overview about existing commercial and non-commercial threat intelligence feeds supporting STIX and TAXII, revealing the commonalities and differences between them. The feeds were evaluated based on their event quality, event timeliness, ease of use, event scope and cost. Some information was not available for all of the evaluated feeds. Such evaluations were marked as N/A. Table 11. shows the scores for each evaluated feed. The scores are a subjective opinions of the thesis's author based on the literature review.

Table 11. Evaluation Summary of different feeds

Service / Feed	Event Quality	Event Timeliness	Ease of use	Event Scope	Cost
Emerging Threats Compromised (Limo)	Good	Good	Excellent	Poor	Excellent
Emerging Threats C&C (Limo)	Excellent	Good	Excellent	Poor	Excellent
Dshield Scanning IPs (Limo)	Poor	Excellent	Excellent	Poor	Excellent
PhishTank (Limo)	Good	Excellent	Excellent	Poor	Excellent
AlienVault OTX	Good	Excellent	Poor	Excellent	Excellent
DHS AIS	Poor	Excellent	Poor	Excellent	Good
IBM X-Force Exchange	Excellent	N/A	Excellent	Excellent	N/A

6.2 Case Study Results

This chapter introduces and summarizes the results of the case study. Data for the case study was collected in a test environment from four STIX/TAXII threat information feeds. The data of the feeds was consumed with STAXX TAXII-client The collected data was exported from the STAXX-client into CSV-files and the indicator analysis was performed in a redundant way by manual comparison and by Excel's built-in conditional formatting for duplicate values. The exported raw data of the Limo and OTX feeds are in Appendix 1. The matched indicators of each of the feeds for each of the target threats are in Appendix 3.

6.2.1 Limo Feeds analysis

Limo feeds introduced a total of 5356 indicators during the test period between 13th and 26th of July 2020. 99,4% of the indicators provided by Limo feeds were received

from the Dshield Scanning IPs feed. Only three different indicator types in total were observed when combining the data of all three feeds as shown in the Figure 10.

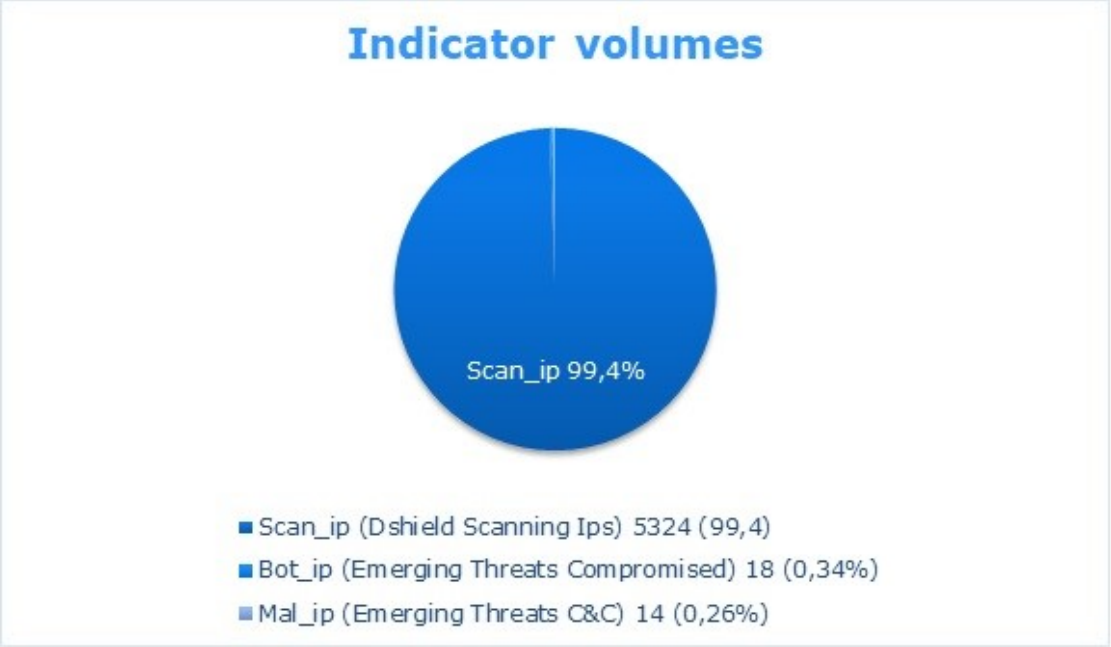

Figure 10. Indicator volumes in Limo feeds

6.2.2 Emerging Threats C&C (Limo)

Event scope

Emerging Threats C&C feed introduced a total of 14 indicators during the test period. Only Indicators of Mal_ip indicator type were observed. A single indicator type results in a poor grade.

Ease of use

The onboarding of this feed was easy. No registration was needed as the feed is accessible with guest user credentials. The feed is STIX/TAXII 2.0 compliant. These two factors result in an excellent grade.

Cost

Setup and usage of this feed was free of charge and the usage did not require extensive human, hardware or software resources, which results in an excellent grade.

Emotet matched indicators

Emerging Threats C&C feed contained six indicators out of 26988 identified Emotet target threat indicators. All the six matched indicators had the Indicator Type Mal_ip. Two of the matched indicators were related to Emotet Epoch 1 command and control servers (C2s) and four of matched indicators were related to Epoch 3 C2s. 99,9% of all the Emotet target threat indicators were missed during the test period. The results are displayed in the Figure 11.

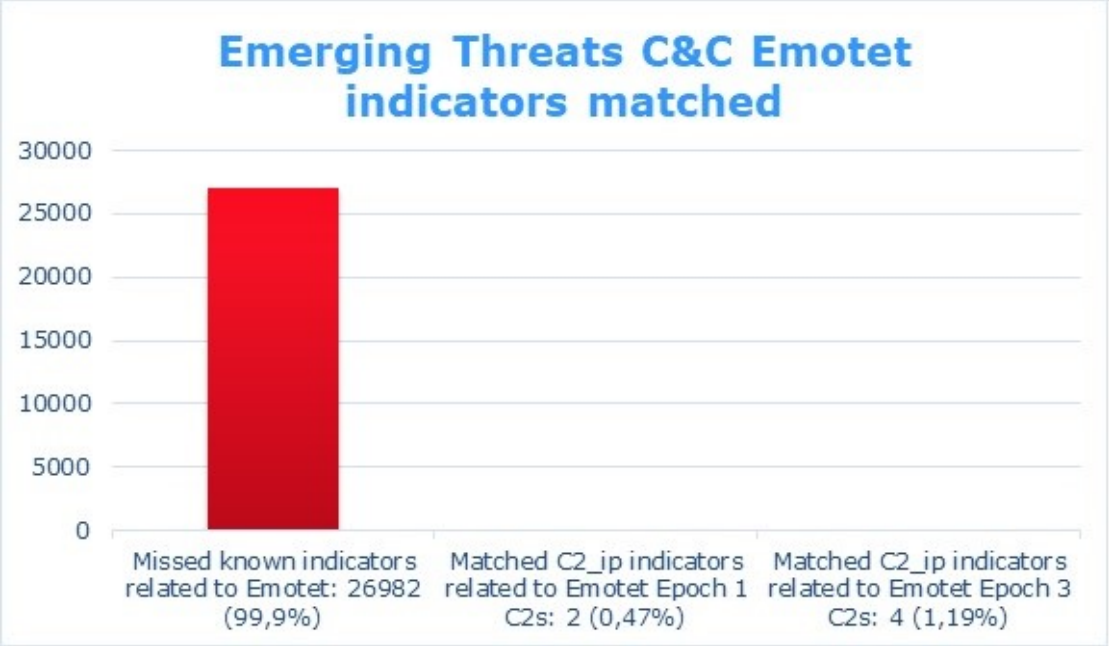

Figure 11. Matched Emotet indicators

Emotet Event timeliness

The indicators which were observed in the Emerging Threats C&C feed had an average time gap of 26 hours between the Cryptolaemus reported time and indicator observation time. As the events in the feed are older than 24 hours, only a poor result for Event timeliness is achieved. Figure 12 displays the variations of indicator timeliness measured from the feed.

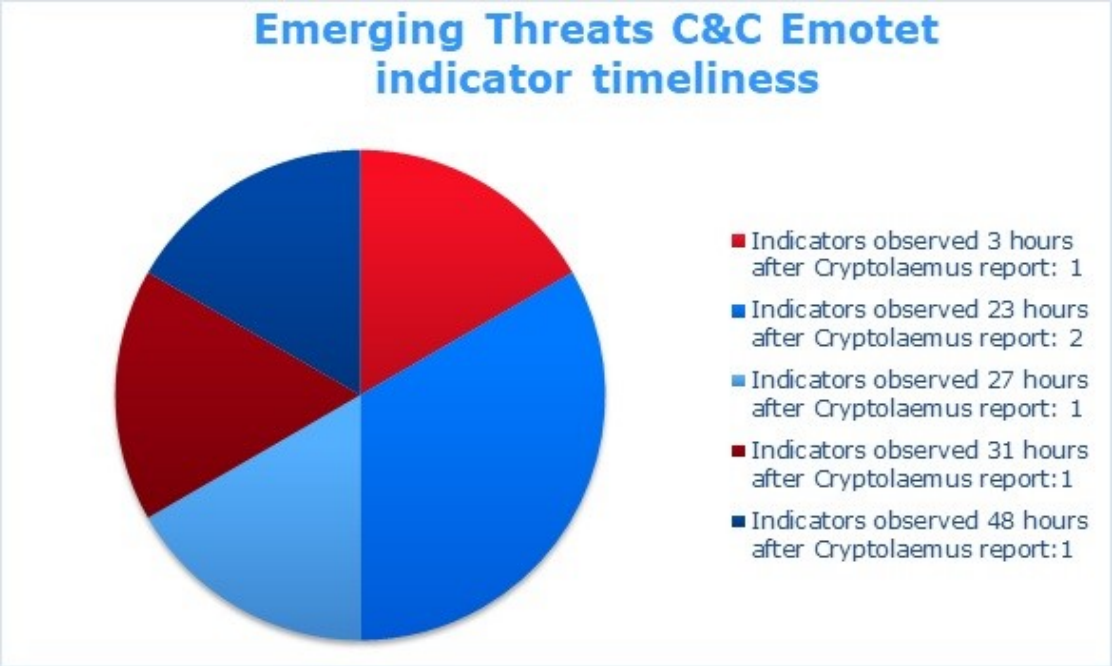

Figure 12. Emotet indicator timeliness

APT29 COVID-19 matched indicators

No IoCs related to APT29 COVID-19 threat were found in this Limo feeds during the test period.

WastedLocker matched indicators

No IoCs related to WastedLocker threat were found in the Limo feeds during the test period.

6.2.3 Emerging Threats Compromised (Limo)

Event scope

Emerging Threats Compromised feed introduced a total of 18 indicators during the test period. Only indicators of Bot_ip indicator type were observed as shown in the Figure 10. A single indicator type results in a poor grade.

Ease of use

The onboarding of this feed was easy. No registration was needed as the feed is accessible with guest user credentials. The feed is STIX/TAXII 2.0 compliant. These two factors result in an excellent grade.

Cost

Setup and usage of this feed was free of charge and the usage did not require extensive human, hardware or software resources, which results in an excellent grade.

Emotet matched indicators

No indicators related to Emotet target threat were found in this feed during the test period.

APT29 COVID-19 matched indicators

No indicators related to APT29 COVID-19 target threat were found in this feed during the test period.

WastedLocker matched indicators

No indicators related to WastedLocker target threat were found in this feed during the test period.

6.2.4 Dshield Scanning IPs (Limo)

Event scope

Dshield Scannin IPs feed introduced a total of 5324 indicators during the test period. Only Indicators of Scan_ip indicator type were observed as shown in the Figure 10. A single indicator type results in a poor grade.

Ease of use

The onboarding of this feed was easy. No registration was needed as the feed is accessible with guest user credentials. The feed is STIX/TAXII 2.0 compliant. These two factors result in an excellent grade.

Cost

Setup and usage of this feed was free of charge and the usage did not require extensive human, hardware or software resources, which results in an excellent grade.

Emotet matched indicators

No indicators related to Emotet target threat were found in this feed during the test period.

APT29 COVID-19 Matched indicators

No indicators related to APT29 COVID-19 target threat were found in this feed during the test period.

WastedLocker Matched indicators

No indicators related to WastedLocker target threat were found in this feed during the test period.

6.2.5 OTX Feed analysis

Event scope

OTX feed provided by AlienVault introduced a total 2292 indicators during the test period. 11 different indicator types were observed in the feed with malicious domain and malicious md5 indicators dominating the volumes as shown in the Figure 13. 11 different indicator types result in an excellent grade.

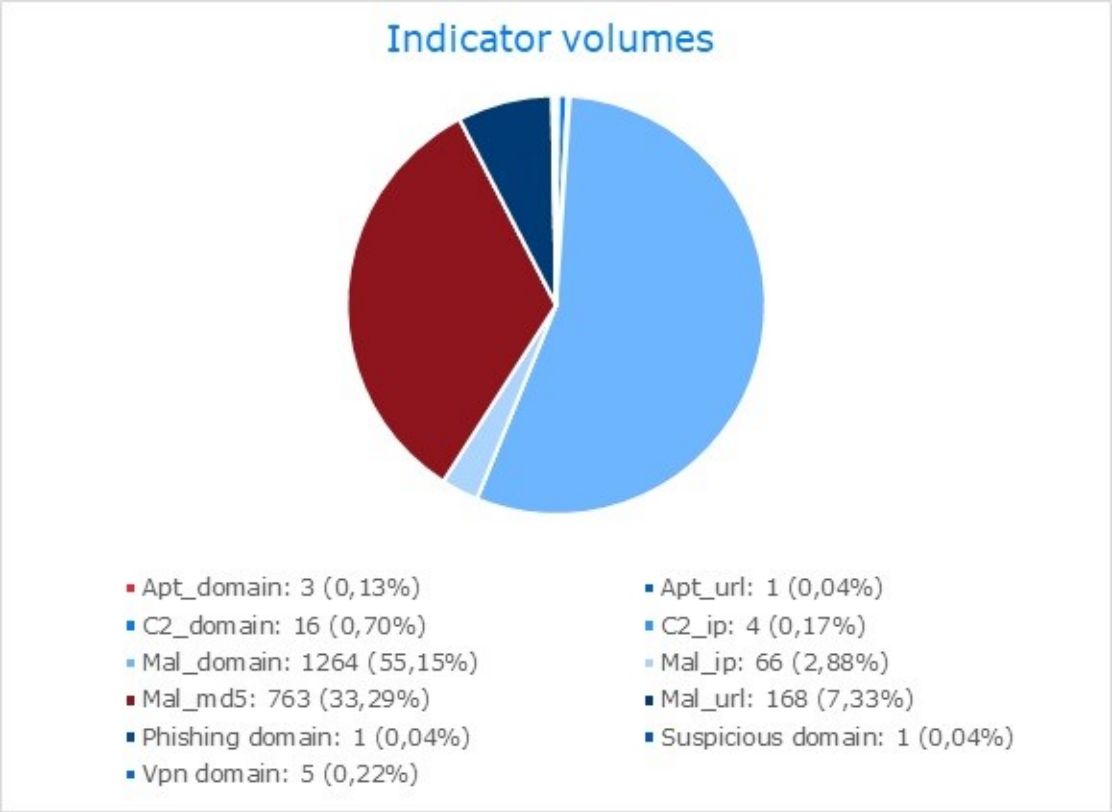

Figure 13. Indicator volumes in OTX feed

Ease of use

The registration process for AlienVault OTX feed was simple and straightforward. The OTX API-key was available after registration and the feed was consumable after the API-key was obtained and entered into the STAXX client. The support for only STIX/TAXII 1.x version results in a poor grade.

Cost

Setup and usage of this feed is free of charge and the usage does not require extensive human, hardware or software resources which results in an excellent grade.

Emotet matched indicators

OTX Feed contained 14 indicators out of 26988 identified indicators related to Emotet target threat. Matched indicator types were mal_ip / C2_ip, Mal_url and Mal_sha256 which were all related to Emotet Epoch 2 indicators as show in the Figure 14.

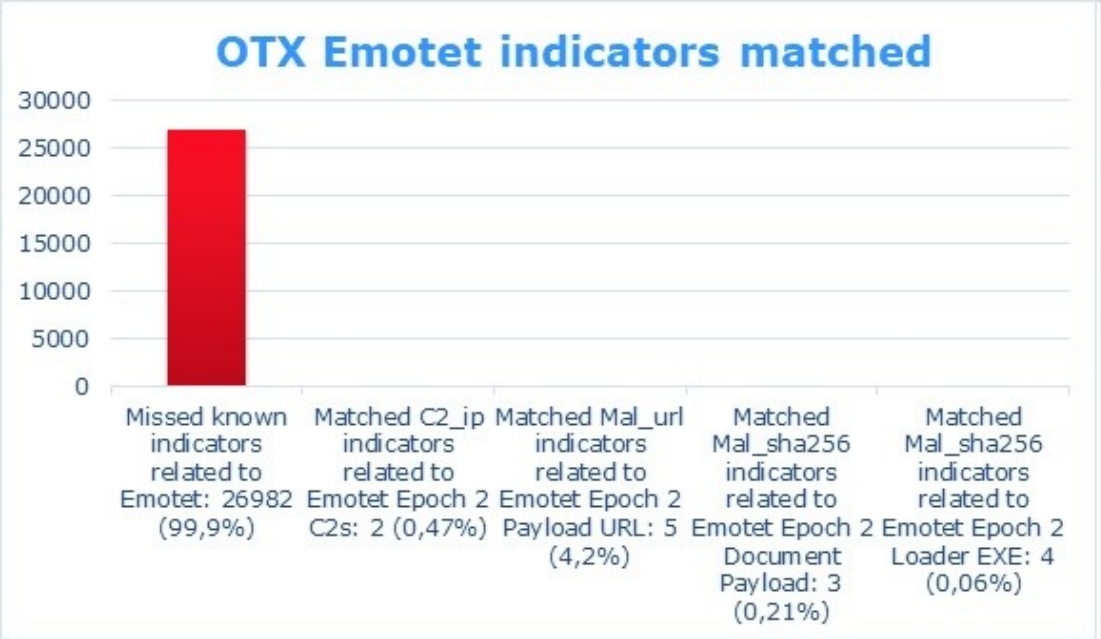

Figure 14. Matched Emotet indicators

Emotet Event timeliness

Matched Mal_url indicators were observed by Cryptolaemus at 17.07.2020 16:04 CEST and reported in the Daily Emotet indicators and Notes for 07/17/20 report at 18.07.2020 07:30 CEST. Matched Mal_url indicators were seen 3 hours and 42 minutes later of the observation time of Cryptolaemus and 12 hours ahead of the reporting time of Cryptolaemus in OTX feed at 17.07.2020 19:46 CEST.

Matched mal_ip / C2_ip indicators were observed and reported by Cryptolaemus at 15.07.2020 01:10 CEST in the Emotet C2 and RSA Key Update - 07/14/2020 23:10 report at 15.07.2020 01:10 CEST. Matched mal_ip / C2_ip indicators were seen 2 days and 7 hours behind the reporting time of Cryptolaemus in OTX feed at 17.07.2020 19:46 CEST.

Matched Mal_sha256 indicators were observed by Cryptolaemus at 17.07.2020 15:04 CEST and 16:04 CEST. The Mal_sha256 indicators were reported by Cryptolaemus in the Daily Emotet IoCs and Notes for 07/17/20 report at 18.07.2020 07:30 CEST. Mal_sha256 indicators were seen 3/4 hours and 42 minutes later of the observation time of Cryptolaemus and 12 hours ahead of the reporting time of Cryptolaemus in OTX feed at 17.07.2020 19:46 CEST. Most of the indicators provided

by OTX feed were 11 hours ahead of the Cryptolaemus reported time. This results in an average time gap of 1,7 hours ahead of Cryptolaemus reported time. A negative value results in an excellent grade. Figure 15 displays the variations of indicator timeliness measured from the feed.

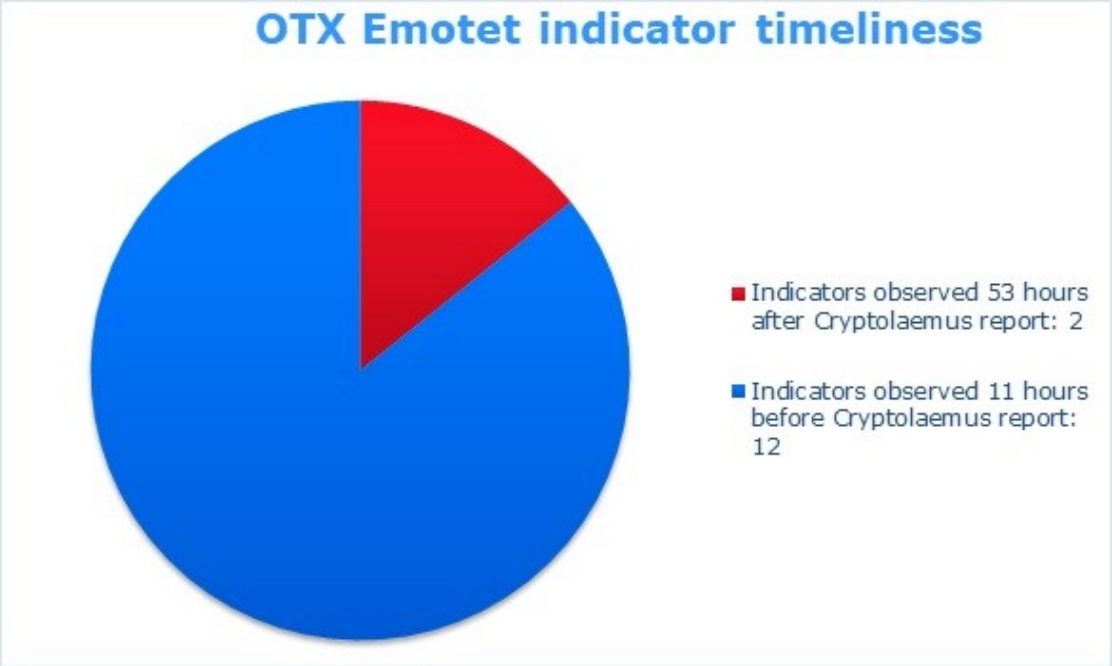

Figure 15. Emotet indicator timeliness

APT29 COVID-19 matched indicators

OTX feed contained 89 indicators out of 93 identified indicators related to APT29 COVID-19 target threat. Only 4 out of 50 Mal_ip indicators were missing from the OTX feed. The missing indicators were all related to WellMess malware as displayed in the Figure 16.

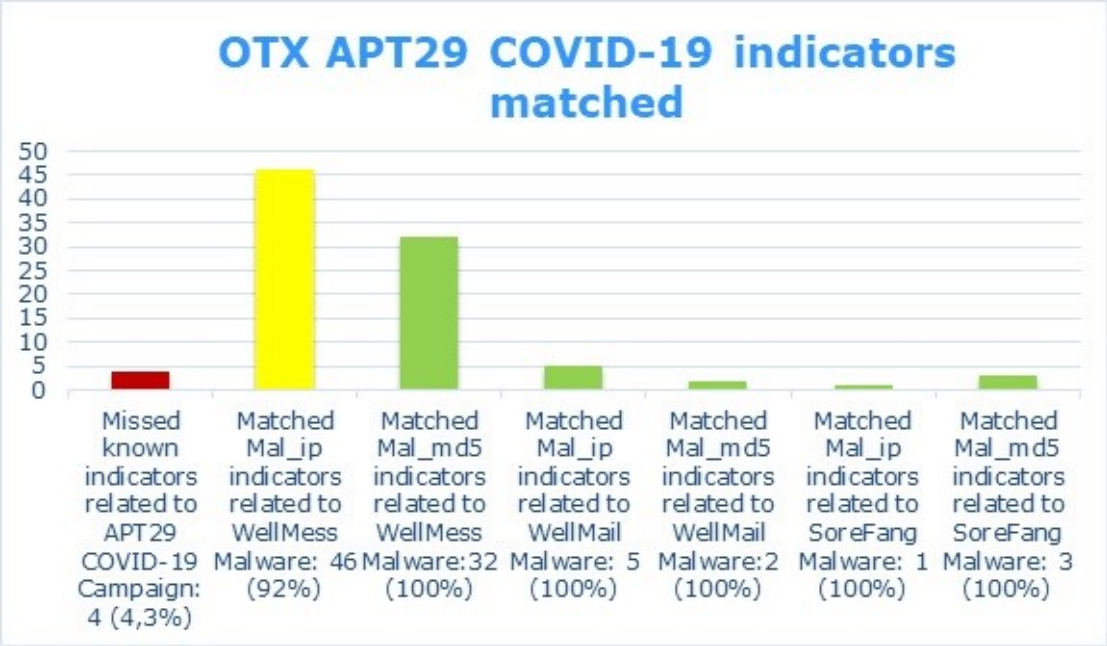

Figure 16. Matched APT29 COVID-19 indicators

APT29 COVID-19 Event timeliness

All the documented indicators related to WellMess, WellMail and SoreFang malware were reported by NCSC at 16.7.2020. The timestamp of the publication is not available, therefore a timestamp of 16.7.2020 00:00 CEST is used. All the matched indicators were observed at 16.7.2020 18:10 CEST in the OTX feed, which is 12 hours after the publication, which results in a good grade.

WastedLocker matched indicators

OTX feed contained eight indicators out of 22 identified indicators related to WastedLocker target threat. The matched indicator types are displayed in the Figure 17.

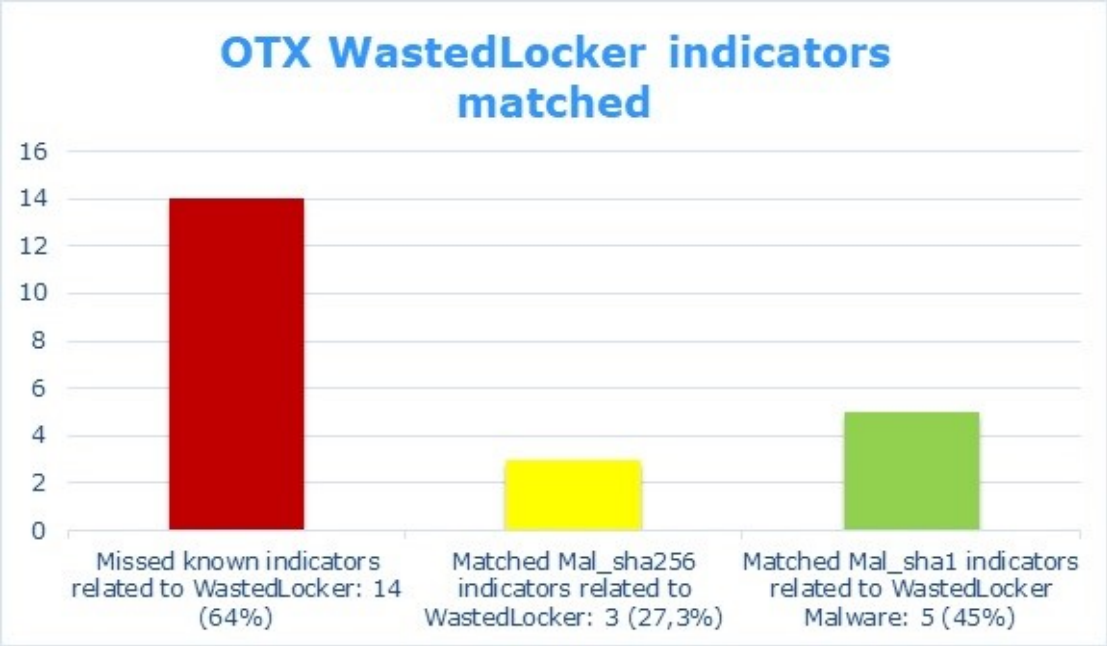

Figure 17. Matched WastedLocker indicators

WastedLocker Event timeliness

All the documented indicators related to WastedLocker target threat were reported by Jim Walter from SentinelOne at 23.7.2020. The timestamp of the publication is not available, therefore a timestamp of 23.7.2020 00:00 CEST is used. All the matched indicators were observed at 24.7.2020 16:26 CEST in the OTX feed, which is 40 hours after the publication. This results in a poor grade.

6.2.6 Summary of case study

The case study defined the measuring points Matched indicators, Event timeliness, Event scope, Ease of use and Cost that were measured in practice by consuming the threat data during the test period. All the measuring points of the literature review except for Event quality were measured and graded with the same evaluation criteria as defined in the literature review. A new measuring point Matched indicators was defined especially for the Case study to measure the volume of matched documented indicators. Table 20. Summarizes the measured results of the case study. Overall percentage of matched indicators value represents the average of matched indicator percentages of all three target threats which for example resulted

in 44% for the AlienVault OTX feed. The Overall Event Timeliness grade represent the averages of timeliness values of all three target threats.

Table 12. Case study summary

Feed	Overall percentage of matched indicators	Overall Event Timeliness	Ease of use	Event Scope	Cost
Emerging Threats C&C (Limo)	0,007%	Poor	Excellent	Poor	Excellent
Emerging Threats Compromised (Limo)	0%	N/A	Excellent	Poor	Excellent
Dshield Scanning IPs (Limo)	0%	N/A	Excellent	Poor	Excellent
AlienVault OTX	44%	Good	Poor	Excellent	Excellent

6.3 Data correlation

While correlating the results of the literature review and the case study we can see that the measured results of the case study support the results of the literature review for Ease of use, Event scope and the Cost. The Event timeliness could not be measured for Emerging Theats Compromised and Dshield Scanning IPs feeds as no matching indicators were found to be able to compare the time stamps. Event timeliness was measured only for Emerging Threats C&C and AlienVault OTX feeds which both scored one grade lower in the case study than in the literature review. The results of the data correlation are illustrated in the Table 13.

Table 13. Data correlation

Feed	Event Timeliness	Ease of use	Event Scope	Cost
Emerging Threats C&C (Limo)	1 grade lower in case study	Equal between both researches	Equal between both researches	Equal between both researches
Emerging Threats Compromised (Limo)	N/A	Equal between both researches	Equal between both researches	Equal between both researches
Dshield Scanning IPs (Limo)	N/A	Equal between both researches	Equal between both researches	Equal between both researches
AlienVault OTX	1 grade lower in case study	Equal between both researches	Equal between both researches	Equal between both researches

7 Evaluation of the results

In this chapter, the research results are examined and presented in more readable format and the research questions are answered.

The first research question "What STIX/TAXII threat information services currently exist?" can be answered by examining the results of the literature review in the Chapter 6.1. The results revealed that currently there is a wide variety of different threat information feeds available supporting STIX format and delivery over TAXII protocol provided by both private and public sectors, both as free to use and as a paid service. The threat data provided by these feeds can be produced by multiple different types of sources such as highly skilled organizations dedicated to tracking a specific threat, sharing communities or honey pots and firewall logs from voluntary users and organizations. The quality of the threat data varies from unfiltered data to very well analyzed and filtered data depending on the provider. The literature review did not cover all the possible STIX/TAXII threat information feeds that currently exists as it was out of scope for this research.

The second research questions "What are the benefits and weaknesses of the existing STIX/TAXII services?" Can be answered likewise by examining the literature review results in the Chapter 6.1. The literature review revealed that some feeds are clearly designed to provide accurate actionable threat information to be used to for example to create automated firewall and IPS rules with minimal need for human intervention. Other feeds on the contrary are designed to provide threat information with possible high rate of false positives to be used for example as supportive threat information to correlate and enrich threat information from other sources with a need for human or artificial intelligence (AI) analysis. Very fast, near real-time event timeliness seems to be achievable with the cost of event quality due to likely increased rate in false positives when there is no time for data verification. Such feeds are aiming to provide threat information as fast as possible to allow the consumers to react on emerging threats in a timely manner when the threat information is still valid and not outdated. A high event quality on the other hand requires data analysis performed by a human analyst, which costs time and therefore reduces the event timeliness of a feed. Some feeds are designed to cover a wide scope of different threat types alone as others have a very narrow scope and need to be used in joint operation with other feeds to promote better situational awareness of an organization.

The results of the case study were mostly aligned with the results of the literature review. The event timeliness scores however were slightly deviant of the literature review results. The deviations in the event timeliness values can partially be explained by possible measuring errors. For example the exact time stamps of the APT29 COVID-19 and WastedLocker target threat reports were not available. The first hour of the day, when the report was published, was used as the time stamp in the case study. This could potentially produce an error of 1 to 23 hours in the measurements. The time stamps of the Emotet indicators were reported sometimes in UTC and sometimes in EDT. Manual extraction of over 20 thousand indicators and manual time-zone conversion is prone for some errors as well. Dshield Scanning IPs threat information feed produced high volumes of data but no Scan_ip indicators were included in the target threat indicators, therefore the data of this feed generated only noise for the data analysis without matching indicators. In a real

environment this feed would probably still be very noisy due to high indicator volumes, but some matching indicators would be expected to be seen as scanning activity in the internet for vulnerable services is quite common and it is also an important part in the first reconnaissance phase of a cyber-attack.

To answer the third research question, "Can consuming of a STIX/TAXII service help to improve situational awareness in Netcloud's Cyber Defense Center and their customer environments?", the answer is yes, but under certain conditions. The question can be answered by examining the results of the case study in the Chapter 6.2. The case study revealed that the Limo feeds provided some indicators related to Emotet Mallware's botnets Epoch 1 and Epoch 3 with an average delay of 26 hours when compared with the Cryptolaemus reports. The coverage of Emotet indicators was widened with the indicators found in the OTX feed. The OTX feed provided some timely accurate indicators specified to Emotet Epoch 2 with an average of 1,7 hours ahead of Cryptolaemus reports. The Indicators to OTX feed must have been sourced from other source than Cryptolaemus reports as they were introduced in the feed before the reports were publisehd by Cryptolaemus. Limo and OTX feeds were able to provide indicators to cover all the three botnets related to Emotet activity. The indicator volumes matched for each Epoch however was very low as only 20 indicators out of total 26988 indicators reported by Cryptolaemus were found in the feeds during the test period. OTX feed was also able to provide timely accurate indicators which covered 96% of the reported indicators related to APT29 COVID-19 campaign. 36% of the reported WastedLocker indicators were seen with a delay of 40 hours in the OTX feed as well.

By consuming just two STIX/TAXII feeds, the consumer would have been able to increase the situational awareness by receiving threat information on almost all reported indicators related to ATP29 COVID-19 campaign and roughly a third of reported malicious hashes related to WastedLocker malware. Additionally some indicators related to all three Emotet botnets would have been available as well. All the indicators would have been available with a maximum delay of 40 hours. Now we get to the part "under certain conditions". Some of the feeds were very voluminous, producing hundreds of indicators daily as displayed in the Figure 18. Such high volumes can generate too much noise and make it very laborsome for an analyst to

identify the relevant data, prioritize, correlate, enrich and make the data actionable by manual work. By consuming STIX/TAXII threat information feeds Netcloud's Cyber Defense Center and any other organization can improve their situational awareness, but the feeds need to be fed into a Threat Intelligence Platform (TIP) to be able to process the threat information into actionable enriched cyber threat intelligence (CTI) more efficiently and possibly automate the process or some parts of it.

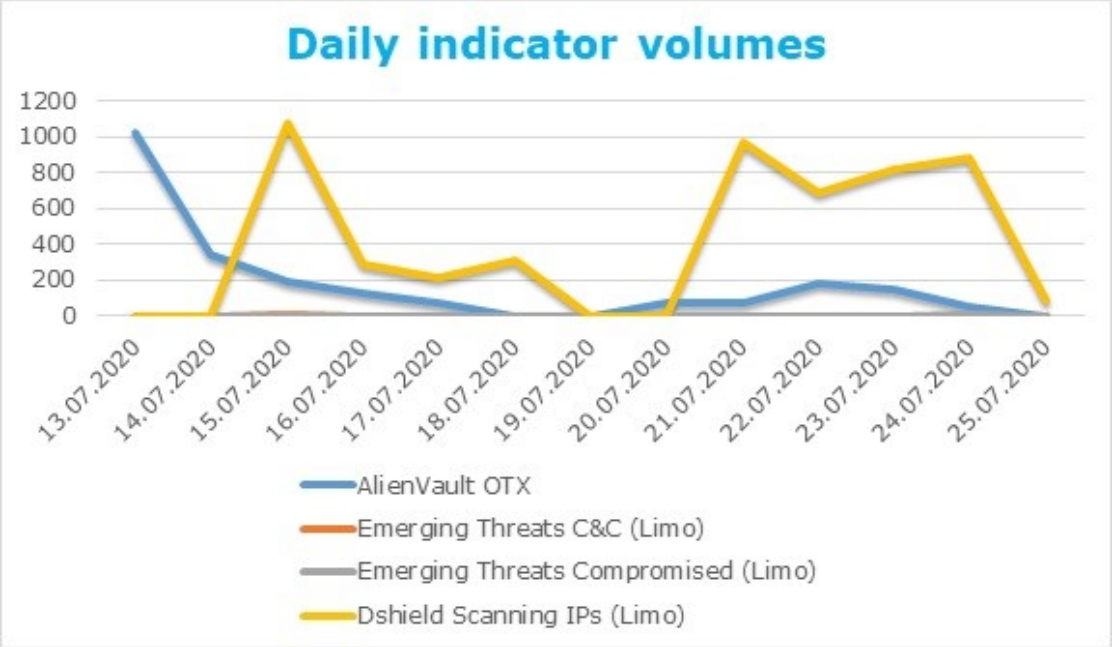

Figure 18. Daily indicator volumes

8 Conclusion

8.1 Discussion

Evaluation of different threat information feeds was needed by Netcloud as such evaluation had not been done in the company before and no previous researches on this matter with easily accessible results were found. The literature review provided easily accessible overview of different STIX/TAXII threat information feeds and their key features which were verified in the case study to some extent. Netcloud can use the results of the literature review and the case study to support their further threat

information feed evaluations to find the threat information feeds and services which serve their business strategy and needs to promote the situational awareness in their Cyber Defense Center and amongst their customers. Netcloud can use the results of the literature review and the case study to choose the suitable feeds after the use case for the feeds has been identified. The evaluation criteria defined in the literature review can also be used to evaluate other threat information feeds which were not covered in this research.

The thesis research process provided a lot of new information on threat information feeds and threat information sharing for the author of the thesis. The literature review was a natural starting point as it provided an overview on the topic along with the necessary ideas and insights to plan and execute the more laborsome and complex parts of the research in the case study. The analysis of the actual threat information data provided by the threat information feeds in the case study really put some flesh on the bones. The very high volume of indicators in the feeds made the data analysis and processing somewhat challenging. The analysis was challenging and time consuming even when the target indicators were already known and well documented. After spending many hours analyzing the threat information a realization was made that a TIP is a necessity when processing high volumes of threat information into actionable threat intelligence in a timely manner.

8.2 Further development

European Union Agency for Cybersecurity (ENISA) recently published a report that identified some of the most pressing actual research topics in the area of cybersecurity around the world. CTI has become an essential part when defending against cyberattacks. Therefore multiple topics around CTI need to be investigated and promoted according to ENISA's report. Existing CTI researches needs to be analyzed and the results should be mapped in a way that they can be compared with existing commercial CTI products and services. The use of open-source CTI could lower the need for CTI skills to easily adopt valuable cyber threat intelligence and therefore the usage of open-CTI should be encouraged. The use of Artificial Intelligence (AI) and other automated tools could reduce the need for manual labor in CTI analysis and promote the effectiveness of CTI related operations (ENISA, 2020).

The report of ENISA reveals that this research is a very actual topic in the area of cybersecurity as it researches CTI. No similar researches were found evaluating STIX/TAXII cyber threat information feeds. The evaluation criteria and measuring points for this research needed to be created from scratch. For the reasons mentioned above a similar research with the same evaluation criteria and measuring points would act as a peer review to validate the reliability of this research and it would also validate if the research methods used in this thesis are adequate.

Some additional ideas for further researches and research questions in the area of CTI were identified during this research process. A research around question, whether consuming additional multiple threat information feeds can be beneficial for better situational awareness and decision making or does the increased noise volume actually negate the benefits of additional feeds, could provide deeper information on how an organization could gain the most benefits out of threat information feeds. This research question could be accompanied with a research covering which advantages does a TIP and AI provide to cyber threat information processing and what would be typical use cases for them.

The research results revealed that some feeds produce raw threat data as fast as possible to achieve excellent timeliness with the cost of data quality and some feeds produce threat data of high quality with the cost of timeliness. A research would be needed to find the best use cases for each of these types of feeds and for the combination of these two.

The STIX threat information feeds covered in this thesis did not use the whole potential of STIX language. Only Indicator and Observable objects were utilized in the feeds. The STIX language could potentially provide many more beneficial objects and concepts which should be adopted to threat information feeds. A potential research in this area could research the use of STIX language more efficiently in threat information feeds.

References

Al-Ibrahim, O., Mohaisen, A., Kamhoua, C., Kwiat, K., Njilla, L. 2017. Beyond Free Riding: Quality of Indicators for Assessing Participation in Information Sharing for Threat Intelligence. Technical Report 2017, University at Buffalo and Air Force Research Lab (88ABW-2017-0416). Accessed 09.10.2020. Retrieved from https://arxiv.org/abs/1702.00552

AT&T. 2020a. About Open Threat Exchange (OTX). Accessed 18.04.2020. Retrieved from https://cybersecurity.att.com/documentation/otx/about-otx.htm

AT&T. 2020b. Welcome to the Alien Labs Open Threat Exchange (OTX). Accessed 18.04.2020. Retrieved from https://cybersecurity.att.com/open-threat-exchange

AT&T. 2019. What's new in OTX. Accessed 18.04.2020. Retrieved from https://cybersecurity.att.com/blogs/labs-research/whats-new-in-otx

Abuse.ch. 2020. About Feodo Tracker. Accessed 18.04.2020. Retrieved from https://feodotracker.abuse.ch/about/

Anomali. 2020. Limo – Free Intel Feed. Accessed 17.04.2020. Retrieved from https://www.anomali.com/community/limo

Anomali. 2017. Anomali Limo - Take the Fast Lane to Threat Intelligence. Accessed 17.04.2020. Retrieved from https://www.anomali.com/blog/anomali-limo-take-the-fast-lane-to-threat-intelligence

Checkpoint. 2020. July's Most Wanted Malware. Accessed 27.09.2020. Retrieved from https://blog.checkpoint.com/2020/08/07/julys-most-wanted-malware-emotet-strikes-again-after-five-month-absence/?web_view=true

Cameron, R. 2015. Mixed Methods Reseach WORKSHOP. Accessed 10.10.2020. Retrieved from https://www.deakin.edu.au/__data/assets/pdf_file/0020/681023/Dr-r-cameron_mixed-methodology.pdf

Cryptolaemus. 2020. Cryptolaemus Pastedump. Accessed 14.10.2020 Retrieved from https://paste.cryptolaemus.com/

Dalziel, H. 2015. How to Define and Build an Effective Cyber Theat Intelligence Capability. Accessed 10.10.2020.
Retrieved from https://janet.finna.fi/, Books24x7 ITPro

DShield. 2020. About Dshield. Accessed 17.04.2020. Retrieved from https://www.dshield.org/about.html

Emerging Threats. 2018. Emerging Threats FAQ. Accessed 17.04.2020. Retrieved from https://doc.emergingthreats.net/bin/view/Main/EmergingFAQ

Emerging Threats. 2017a. Known Compromised Hosts. Accessed 18.04.2020. Retrieved from https://doc.emergingthreats.net/bin/view/Main/CompromisedHost

Emerging Threats. 2017b. Known Bot Command and Control Rules. Accessed 18.04.2020. Retrieved from https://doc.emergingthreats.net/bin/view/Main/BotCC

ENISA. 2020. ENISA Thread Landscape. Accessed 23.10.2020. Retrieved from https://www.enisa.europa.eu/publications/enisa-threat-landscape-2020-research-topics

Garrido-Pelaz, R., González-Manzano, l., Pastrana, S. 2016. Shall we collaborate? A model to analyse the benefits of information sharing. Proceedings of the 2016 ACM on Workshop on Information Sharing and Collaborative Security Pages 15-24. Accessed 09.10.2020. Retrieved from https://arxiv.org/abs/1607.08774v1

Homeland Seurity. 2015. DHS Leads Effort to Transition Automated Cybersecurity Information Sharing Specifications to International Community. Accessed 16.03.2020. Retrieved from https://www.dhs.gov/blog/2015/07/23/dhs-leads-effort-transition-automated-cybersecurity-information-sharing

Homeland Security. 2020. Automated Indicator Sharing (AIS). Accessed 16.03.2020. Retrieved from https://www.us-cert.gov/ais

HSSEDI. N.d. Analyzing and Sharing Cyber Threat Intelligence. Accessed 18.03.2020. Retrieved from https://www.first.org/resources/papers/munich2016/wunder-stix-taxii-Overview.pdf

IBM. 2020a. IBM X-Force Exchange. Accessed 19.04.2020. Retrieved from https://www.ibm.com/fi-en/marketplace/ibm-xforce-exchange/details

IBM. 2020b. IBM X-Force Exchange API Documentation. Accessed 19.04.2020. Retrieved from https://api.xforce.ibmcloud.com/doc/

IBM. 2020c. IBM X-Force FAQ. Accessed 19.04.2020. Retrieved from https://www.ibm.com/products/ibm-xforce-exchange/faq

Johnson, C., Badger, L., Waltermire, D., Snyder J., Skorupka, C. 2016. Guide to Cyber Threat Information Sharing. National Insitute of Standards and Technology (NIST) Special Publication 800-150 Accessed on 6. February 2020. Retrieved from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf

Kananen, J. 2017. Laadullinen tutkimus pro graduna ja opinnäytetyönä. Accessed 10.10.2020. Retrieved from https://janet.finna.fi/, Booky.fi

Kaspersky. 2020. Analysis of WastedLocker targeted ransomware. Accessed 14.10.2020. Retrieved from https://www.kaspersky.com/blog/wastedlocker-garmin-incident/36626/

Kokkonen, T. 2016. Anomaly-Based Online Intrusion Detection System as a Sensor for Cyber Security Situational Awareness System. Doctoral Dissertation. Faculty of Information Tehcnology. Universtity of Jyväskylä. Accessed 15.04.2020. Retrieved from http://urn.fi/URN:ISBN:978-951-39-6832-8

Lauf, F., Kuziemsky, C. 2017. Handbook of eHealth Evaluation: An Evidence-based Approach. Accessed 10.10.2020. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK481583/

Mitre. 2020. Situation Awareness. Accessed 09.10.2020. Retrieved from https://www.mitre.org/capabilities/cybersecurity/situation-awareness

NCSC. 2020. Advisory: APT29 targets COVID-19 vaccine development. Accessed 27.09.2020
Retrieved from https://www.ncsc.gov.uk/files/Advisory-APT29-targets-COVID-19-vaccine-development.pdf

Nist. 2016. NIST Special Publication 800-150. Accessed 17.03.2020. Retrieved from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf

Oasis CTI. 2016a. STIX™ Version 1.2.1. Part 1: Overview. Accessed 16.03.2020. Retrieved from http://docs.oasis-open.org/cti/stix/v1.2.1/stix-v1.2.1-part1-overview.html

Oasis CTI. 2017. STIX™ Version 2.0. Part 1: STIX Core Concepts. Accessed 16.03.2020. Retrieved from http://docs.oasis-open.org/cti/stix/v2.0/cs01/part1-stix-core/stix-v2.0-cs01-part1-stix-core.html

Oasis Open. N.d. Accessed 21.03.2020. Retrieved from https://oasis-open.github.io/cti-documentation/stix/compare.html

Oasis CTI. 2016b. TAXII™ Version 1.1.1. Part 1: Overview. Edited by Mark Davidson, Charles Schmidt, and Bret Jordan. 05 May 2016. OASIS Committee Specification 01. Accessed 21.03.2020. Retrieved from http://docs.oasis-open.org/cti/taxii/v1.1.1/cs01/part1-overview/taxii-v1.1.1-cs01-part1-overview.html#_Toc450734054

Oasis CTI. 2018. TAXII™ Version 2.1. Edited by Bret Jordan and Drew Varner. 15 December 2018. OASIS Committee Specification Draft 02 / Public Review Draft 01. Accessed 21.03.2020. Retrieved from https://docs.oasis-open.org/cti/taxii/v2.1/csprd01/taxii-v2.1-csprd01.html

Oasis CTI. 2020. STIX™ Version 2.1. Accesssed 06.10.2020. Retrieved from https://docs.oasis-open.org/cti/stix/v2.1/cs01/stix-v2.1-cs01.html#_axjijf603msy

Phishtank. 2020a. Phistank FAQ. Accessed 17.04.2020. Retrieved from https://www.phishtank.com/faq.php#whatisphishtank

Phishtank. 2020b. Phistank API Information. Accessed 17.04.2020. Retrieved from https://www.phishtank.com/api_info.php

Phishtank. 2020c. Phistank API Information. Accessed 17.04.2020. Retrieved from https://www.phishtank.com/developer_info.php

Shadowserver. 2020. What we do. Accessed 18.04.2020. Retrieved from https://www.shadowserver.org/what-we-do/

Simons, H. 2009. Case study research in practice. Accessed 10.10.2020. Retrieved from https://janet.finna.fi/, Ebook Central Academic Complete International Edition

Traficom. 2020. Emotet-haittaohlemaa levitetään aktiivisesti Suomessa. Accessed 28.09.2020. Retrieved from https://www.kyberturvallisuuskeskus.fi/fi/emotet-haittaohjelmaa-levitetaan-aktiivisesti-suomessa

Walter. J. 2020. SentinelOne, WastedLocker Ransomware: Abusing ADS and NTFS File Attributes. Accessed 14.10.2020. Retrieved from https://labs.sentinelone.com/wastedlocker-ransomware-abusing-ads-and-ntfs-file-attributes/

Appendices

Appendix 1. Exported raw data from OTX and Limo feeds

Exported raw data of OTX and Limo feeds during the test period is available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/Raw_Data_Exports_OTX_and_LIMO_13-26.7.2020.xlsx

Appendix 2. Extracted target threat indicators

Extracted indicators for target threat Emotet available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/Extracted_Emotet_Target_Threat_IoCs.xlsx

Extracted indicators for target threat ATP29 available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/Extracted_ATP29_Target_Threat_IoCs.xlsx

Extracted indicators for target threat WastedLocker available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/Extracted_WastedLocker_Target_Threat_IoCs.xlsx

Appendix 3. Matched indicators

Mached Emotet Indicators for Limo feeds available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/Limo__Matched_Emotet_IoCs.xlsx

Mached Emotet Indicators for OTX feed available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/OTX__Matched_Emotet_IoCs.xlsx

Mached ATP29 Indicators for OTX feed available here:
https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/OTX__Matched_ATP29_IoCs.xlsx

Mached WastedLocker Indicators for OTX feed available here:

https://github.com/TaalaJ/Evaluation-of-Threat-Information-Feeds-for-a-Cyber-Defense-Center/blob/main/OTX__Matched_WastedLocker_IoCs.xlsx