TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science

ITC70LT

Kim Vahturov 163172IVCM

USING INDICATORS OF COMPROMISE TO AUTOMATE INCIDENT TRIAGE. PROOF OF CONCEPT.

Master's thesis

Supervisor: Toomas Lepik
Master of Science
TUT Early Stage Researcher
Co-supervisor: Lauri Palkmets
Master of Science
NATO Incident Investigation Officer

Tallinn 2018

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

ITC70LT

Kim Vahturov 163172IVCM

KOMPROMITEERIMISE INDIKAATORITE KASUTAMINE KÜBERINTSIDENTIDE TRIAAŽI AUTOMATISEERIMISEKS. KONTSEPTSIOONI TÕENDUS.

Magistritöö

Juhendaja: Toomas Lepik
Magister
Nooremteadur
Kaasuhendaja: Lauri Palkmets
Magister
NATO intsidendihaldur

Tallinn 2018

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Kim Vahturov.

23.04.2018

Abstract

Effective protection of information systems implies availability of financial investments, automation of security-related workflows and knowledge about recent tactics, techniques and procedures used by system attackers. This research aims to develop an open-source tool framework, which would enhance incident management by using indicators of compromise for automated detection and triage of cyber incidents.

First part of this thesis represents a theoretical study of incident management best-practises and threat intelligence topics, focusing on analysis of requirements for effective implementation of a potential solution. Second part of thesis describes a case study, which was conducted during practical part of this research to test efficiency of the proposed tool framework.

Research confirmed that implementation of the framework allows to automate disclosure of malicious behavior within a given environment, triage and register determined incidents in an automated manner to support further incident handling. Proposed tool framework can be implemented at no cost, allow security teams to increase incident detection rates and partially automate security-related workflows.

This thesis is written in English and is 64 pages long, including 5 chapters, 22 figures and 3 tables.

Annotatsioon

Kompromiteerimise indikaatorite kasutamine küberintsidentide triaaži automatiseerimiseks. Kontseptsiooni tõendus.

Efektiivse infosüsteemide kaitse korraldamise eelduseks on rahaliste vahendite olemasolu, turvalisuse tagamisega seotud protsesside automatiseerimine ja arusaam süsteemi ründajate poolt kasutatavatest taktikast, tehnikast ja protseduuridest. Käesoleva uurimistöö eesmärk on töötada välja raamistik avatud lähtekoodiga tarkvaradest, mis tõhustaks intsidendihaldust läbi küberintsidentide tuvastamise ja nende triaaži teostamise automatiseerimise, kasutades kompromiteerimise indikaatoreid..

Töö esimeses osas analüüsib autor teoreetilist kirjandust intsidentide halduse ja ohuteadmuse teemadel. Selle raames keskendub autor potentsiaalse lahenduse kriteeriumite väljatöötamisele. Töö teises osas kirjeldab autor praktiliselt teostatud juhtumiuuringut, mille raames testis tema pakutut lahenduse effektiivsust.

Uurimistöö tõestas, et pakutud lahenduse rakendamine võimaldab automatiseerida pahaloomulise tegevuse tuvastamise, tuvastatud intsidentide triaaži teostamise ja nende registreerimise edasise menetlemise hõlbustamiseks. Autori poolt pakutud lahendus võimaldab tõsta intsidentide tuvastamise effektiivsust ja osaliselt automatiseerida intsidentide haldamisega seotud protsesse ilma täiendavate eelarveliste vahendite kasutamiseta.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 64 leheküljel, 5 peatükki, 22 joonist, 3 tabelit.

List of abbreviations and terms

SNMP	Simple Network Management Protocol
CSIRT	Computer security incident response team
SOC	Security operation center
SIEM	Security event and information management system
IOC	Indicator of compromise
SLA	Service-level agreement
IT	Information technology
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
NTP	Network Time Protocol
TTP	Tactics, Techniques and Procedures
ENISA	European Union Agency for Network and Information Security
EU	European Union
ITIL	Information Technology Infrastructure Library
NIST	US National Institute of Standards and Technology
SANS	SANS Institute
CVSS	The Common Vulnerability Scoring System
DNS	Domain Name System
DMZ	Demilitarized zone
NATO	The North Atlantic Treaty Organization
TI	Threat Intelligence
CEO	Chief executive officer
CFO	Chief financial officer
CIO	Chief information officer
ISAC	Information Sharing and Analysis Center
C&C	Command and control
API	Application programming interface
LAN	Local area network

Table of contents

List of figures

List of tables

1 Introduction

1.1 Background.

As a result of computer technology evolution, computers have been grouped into networks and networks connected with each other, making up a global network of interconnected devices. This has become possible due to implementation of standards for network communication [1]. Open standards allow technology vendors to develop their products in such a way, that complex processes of interconnecting network hosts occur with minimal human intervention.

In 1988 approximately 5% of all computers, connected via the Internet have been infected with the Morris worm, which abused open standards of SMTP protocol to propagate itself across the global network [2]. This cyber incident had several significant consequences, including establishment of the first CSIRT. Since that time CSIRTs have evolved from a groups of system administrators to highly trained and efficiently organized cyber security and digital forensic professionals [3]. Computer systems have concurrently become more complex and their complexity created advanced challenges in the field of cyber security. On the other hand, while the creator of the Morris worm was inspired by curiosity, modern illegal cyber activities are commonly associated with totally different motives. Cybercrime has escalated to an organized level where actors are often motivated with financial profit or even political interests [4].

Cyber incident response capabilities need to adequately correspond to the sophistication degree of modern cyber-attacks. On a cyber battlefield CSIRTs rely on defense systems, which analyze security events originating from diverse sources like perimeter defense systems, hosts, applications and network sensors. Devices and various software components are used to monitor the security events, hunt the indicators of possible attacks and exchange threat-related information. The highest maturity level CSIRTs contribute to system security through investments into offensive self-defense capabilities. However, the effectiveness of investments into advanced cyber domains can be significantly reduced, if contribution of lower-level security implementations that create foundation of

defense capabilities is insufficient [5]. Regardless of resources dedicated to a particular CSIRTs, most of them are still victims of the so-called "Fortification principle". It states, that the attackers of information systems stay in a beneficial position because [6]:

- they only need to abuse a single system vulnerability, while CSIRTs have to guard them all;
- CSIRTs are too slow to determine and eliminate system vulnerabilities before they could be exploited by the attackers;
- To effectively defend the systems, CSIRTs need significantly more resources in comparison to adversaries who try to find a way in.

Michael Walker, the Darpa cybersecurity program manager suggested that automation of cyber security processes and enhancing computers' sleuthing capabilities should allow to overcome "Fortification principle" and place system defenders in advantageous position [6].

Another important challenge in cyber domain is significant increase in amount of security-related data, which needs to be processed by CSIRTs. In 2017 a "Sick SOC" term has been proposed by William Cole, a VP and Global CTO at FireEye company. The term describes SOCs in context of their maturity as reactive alert-response teams, who mostly stop legacy threats that bounce around the Internet (see figure 1) [7]. Instead of facing actual problems of their security systems, "Sick SOCs" mainly attempt to collect as much alerts as possible, and as a consequence, get overloaded by the amount of alerts aggregated from different sources.

The number of triggered events depends on the amount of particular organization's employees, network devices and their complexity. In big enterprises it may reach 100 billion to 1 trillion on a daily basis, which makes it impossible for SOC personnel to take every single event into account. Overwhelming volumes on information combine benign events together with events triggered by malicious activity, which creates a necessity of distinguishing between different types of alarms [8].

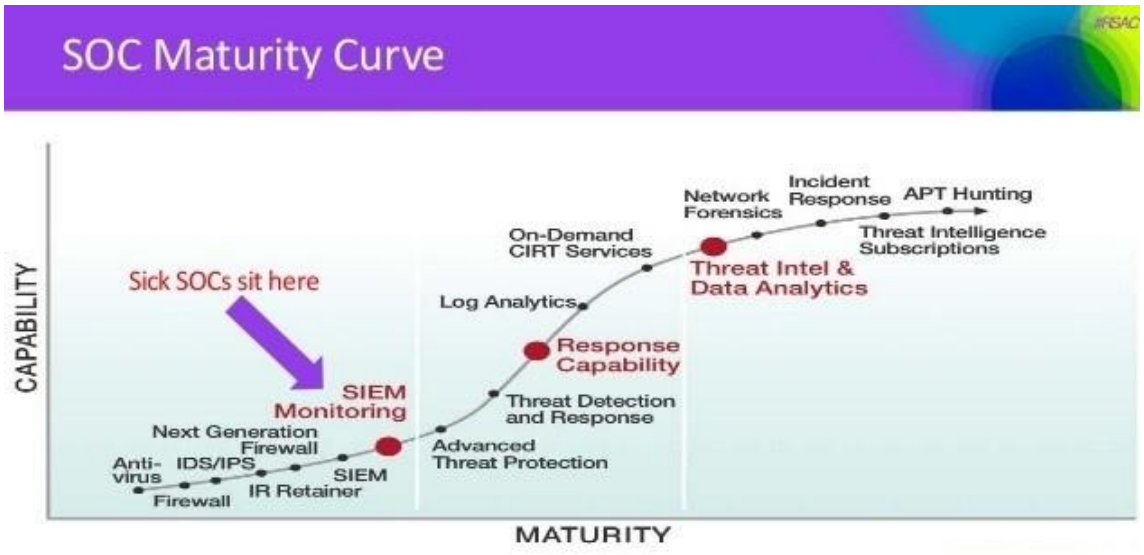


Figure 1. SOC maturity curve. Source: [7]

Some SOCs claim that 10 minutes is an acceptable time for filtering cyber incidents from the other events provided by SIEMs, but they would desire to reduce this time to 1 minute, or even less [8]. The process of validating cyber incidents through analysis of aggregated security events is a part of another process, known as cyber incident triage.

Cyber incident triage is a process, which is used to support decision-making by assigning priorities to incidents for effective usage of organizational resources. Triage helps to determine severity of analyzed events and manage available resources for the most effective response, and mitigation of negative consequences of a particular incident. Development of SIEMs has had a very significant meaning for security experts, because it enables automatic correlation of such amount of data that can't be interpreted by a human intelligence anymore. SIEM is a good example of successful cyber security processes automation, however the triage of SIEM's output and security-related data originating from other sources is still often performed by SOC personnel manually.

Automating triage process would facilitate mitigation of incidents' impact, because faster incident triage results in faster response activities. Correct selection of suitable tools and methods used for triage purposes may provide possibilities to accelerate analysis of security events, and even help to achieve proactive defense objectives. Faster triage would assist "Sick SOCs" to overcome the problem of analyzing overwhelming amounts of data, and thus move onto the next maturity tier. However, organizations are typically required to make investments into specific security tools or expert intelligence to experience the benefits of automated solutions. Since CSIRTs are usually limited in terms of budget or

personnel trainings, it is not uncommon that suitable tools and/or expertise for a particular organization become unavailable [9].

1.2 Scope of the thesis.

The purpose of this research is to find an answer to the question: "is it possible to automate cyber incident triage process by using free software and publicly available information?" Such solution would allow organizations to increase their cyber incident response capabilities without the need of additional financial investments. Automation of workflows would allow cyber security teams to move onto higher maturity tier by redirecting human intelligence towards advanced cyber security domains, increasing CSIRT's overall efficiency [7].

Cyber incidents can be detected through many different means. Abnormal behavior can be reported by organization's personnel, or detected automatically with the help of network and host monitoring tools [10], [11]. In other cases an organization can be contacted by external party (e.g. Internet service provider) that observes malicious activities associated with organization's infrastructure. However, this research is specifically focused on proposing solution for triaging incidents, which can be determined with help of network-based IOC check within a given environment.

Author will try to answer the main research question by testing capabilities of free and open source software to implement IOCs for detecting and triaging simulated incidents in a prepared virtual environment.

1.3 Research methods and thesis outline.

The following steps are taken by author to answer to the formulated question:

Conduct a research of existing literature on cyber incident triage topic.

The initial step is to analyze existing methodologies of cyber incident triage and its dependencies with other incident management sub processes. The guidelines and best practices related to incident management and particularly triage process are analyzed in Chapter 2. The chapter also includes investigation of cyber threat assessment

methodologies and factors that should be considered during estimating severity of cyber incidents.

Perform a theoretical study on threat intelligence sharing topic.

Benefits and challenges of sharing cyber threat information are discussed in Chapter 3. Author analyzes which methods should be used to overcome challenges related to threat intelligence consumption, how to evaluate quality of received information and perform threat hunting within a particular environment using indicators of compromise.

Design a framework for partial automating of incident triage.

In Chapter 4 author analyzes existing software products and their suitability for supporting incident triage automation. Analysis is performed considering the requirements determined during theoretical research and particular software interoperability capabilities.

Author conducted a case study by simulating malicious activities in a virtual test environment to test the efficiency of designed solution. Chapter 4 describes case study process, highlights deficiencies of suggested framework, and proposes solutions to mitigate the shortcomings.

Conclusion.

Chapter 5 summarizes the thesis by evaluating if investigation methods and results produced by case studies answered the main question of the research. Author also proposes recommendations for future research to improve the results of automating incident triage process and enhance capabilities of the proposed tool framework.

2 Incident management and triage processes.

2.1 Triage as a sub process of incident management

Historically activities performed by CSIRTs have been defined by the terms “incident handling” and “incident response”, and they were reactive in their nature. Actions were taken to resolve or mitigate an incident after it has been detected. As CSIRTs evolved, the set of their activities also expanded; preventing incidents from happening through securing and hardening infrastructure, conducting trainings, active infrastructure monitoring and scanning, as well as sharing incident-related data, has become a significant challenge. The term “incident management” includes various services and functions that may be performed by organization, and it should not be perceived as simply responding to an incident when it happens. This process includes more than application of security technologies, it rather represents a multilayered strategy consisting of technical and organizational approaches for preventing and mitigating cyber incidents. [12].

Effective incident management requires a developed plan of action, which should be integrated into existing business processes and organizational structures. A good incident management plan should protect and secure critical business assets, functions and processes from internal and external threats. Such plan should be dynamic in terms of its ability to adapt to changing business requirements ant threat landscape [13]. For this reason it should be continuously updated following a life cycle approach, and particularly describe the following processes, which occur in sequence and are build one upon another [10], [12] - [14]:

- prepare/sustain/improve – planning, implementing, sustaining and enhancing incident management capability
- protect infrastructure – implementation of infrastructure protection changes and improvements, conducting proactive scanning and monitoring, performing security and risk evaluations

- detect events - receiving and reviewing event information and incident reports, analyzing alerts and indicators
- triage events – categorization and correlation of security events, determining and prioritizing incidents
- respond – detailed analysis of incidents; planning, coordinating and implementing containment, mitigation, resolution and recovery strategy
- lessons learned - summarizing and documenting incident-related information, which may be beneficial in terms of preventing future incidents by improving previous sub processes based on gathered knowledge.

Preparation and protection in context of incident management are continuous ongoing processes that involve establishing policies and procedures, conducting risk assessment, managing personnel resources, technologies and infrastructure. Effectiveness of incident management activities and other processes highly depend on how effectively are implemented preparation processes and protection capabilities [12].

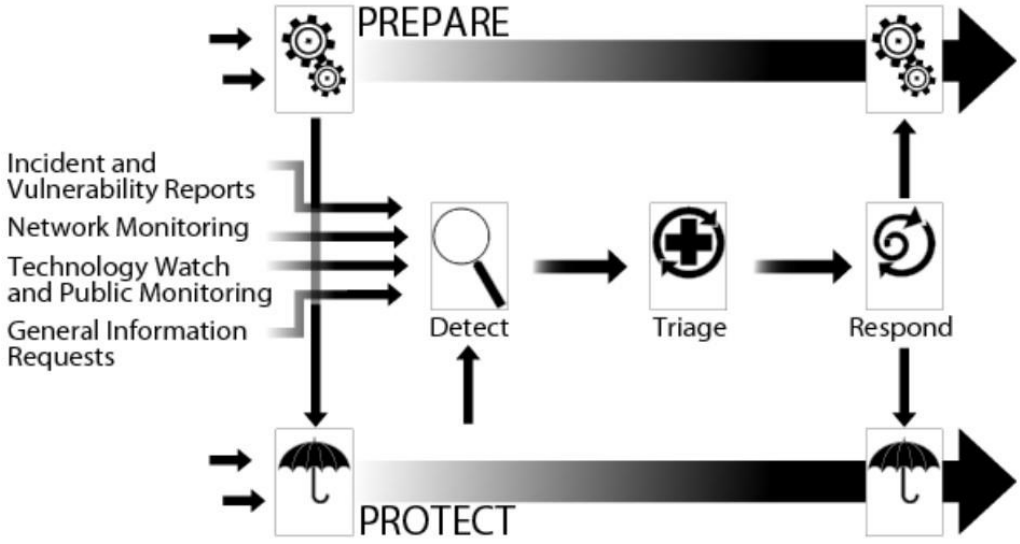


Figure 2. High-level incident management processes. Source: [12].

Established policies are important in context of incident management, since policies provide CSIRTs a legal framework for identifying incidents through written sets of principles, rules and procedures within a particular organization. Personnel resources management is significant for providing CSIRTs professional capabilities and necessary

trainings to ensure that CSIRT employees know how to perform their duties. Additionally, personnel resource management plans can typically describe rules of communication with specific individuals during investigation of an event. Incident response may be delayed, if CSIRT contacts wrong people, or lacks know-how to mitigate the incident using its internal employees [10], [14]. Preparations for an effective incident handling also include providing CSIRT with necessary tools, including any available software and hardware that can be useful, and appropriate system access permissions to perform their job [12], [14].

Protection process relates to implementing changes in infrastructure and existing configurations, which may be dictated by results of risk assessment, security audits, incident analysis, vulnerability scans, or processing threat-related information. Infrastructure changes aim to improve organization's environment by modifying system configurations, patching discovered vulnerabilities, implementing and improving threat mitigation strategies and best practices [12].

The "detect", "triage" and "respond" processes depend on each other in a way, where output of previous process is used as an input for the following. Incident and vulnerability reports, information requests, or suspicious event-related data can be gathered either reactively, or proactively [12].

- Reactive detection mainly occurs when either internal (e.g. an employee of the organization) or external (e.g. another CSIRT) parties report abnormal or malicious activities, related to CSIRTs constituency, or organization's infrastructure [12].

- Proactive detection assumes that CSIRT performs monitoring of a variety of data (e.g. different logs, netflow data, etc.), actively searches for threat-related information, or uses external services to detect indicators of malicious activity within organization's environment [12].

In both cases, the detected activity or information is passed on to the triage process as a report, alert, or other kind of notification. To support effective triage, passed information should include as much relevant data as possible, e.g. incident or vulnerability reports, indicators of suspicious system or network behavior and results of preliminary data analysis, if it has been performed before handing off the information [12].

The triage should determine, if deviations from normal network and system behavior can be considered incidents and in such cases assigned for handling or response. The process itself is described in more details in the following section. However, in context on understanding incident management workflow it is important to describe “triage” and “respond” processes dependencies.

In some cases triage verifies events, and associated information that is passed from “detect” process, as incidents. In such cases all incident-relevant information is enriched and forwarded to “respond” process for a more detailed incident analysis, mitigation and response. Triage enriches the information by describing initial class and priority of verified incidents and their possible relations with other incidents, which have occurred in the past or occur currently. In other cases security events may be forwarded to “respond” process for additional analysis, before they can be verified as incidents, or classified and prioritized appropriately. In both cases triage assigns the responsible incident handler, who processes forwarded data during “respond” process [12].

Incident response includes implementing appropriate technical, management and legal countermeasures to resolve or mitigate incidents. For the most effective response these countermeasures should be implemented in a coordinated manner, which assumes that information sharing and communication channels are established between members of organization as well as with external parties [10], [12], [13].

Summarizing and documenting incidents that have been resolved or mitigated, provides a possibility to evaluate existing processes and determine problems in policies, infrastructure as well as employees training and awareness. Lessons learned are passed to the “prepare” process for system hardening and future incidents containment through organizational or technical changes. This closes incident lifecycle in terms of described incident management process [12].

2.2 Incident triage sub processes and methodologies.

Triage is an essential process of incident management and a prerequisite for effective incident handling. It represents initial analysis of security events, which is performed through categorizing, correlating, prioritizing, and assigning incoming information to appropriate handlers [12], [15]. Such information may include:

- Security events
- Incident reports
- Vulnerability reports
- Information requests
- Threat intelligence feeds

Effective implementation of incident triage assumes that such information is passed through a single point of contact, regardless of the information delivery means (e.g. e-mail, telephone, etc.). This is important because it limits the ability of constituents and others to bypass triage process, and enables systematic information redistribution and handling within a CSIRT [16]. It is also important in terms of providing a single repository for tracking status of reported events or activities, which helps to identify potential security problems within organization's environment and prioritize workload of employees [12].

Triage begins with categorization and correlation of received information. Once the information is passed to the triage process it is initially sorted based on the information content (categorized) and compared with the data in available repositories regarding security events and incidents (correlated) [16]. Sorting and correlating received data helps to verify security incidents through determining possible interconnections between host and network activities. A security event can be verified once CSIRT determines affected networks, system, or applications (attack targets), and attack methods or actions (see figure 3) [15]. Received data is additionally correlated with information that is related to current or past incidents, to determine their possible dependencies, or verify the data as new incident. Identifying unique incidents, as well as determining interconnections of multiple incidents can support further correlation of events by enriching incident description with additional information fields (tags) [16].

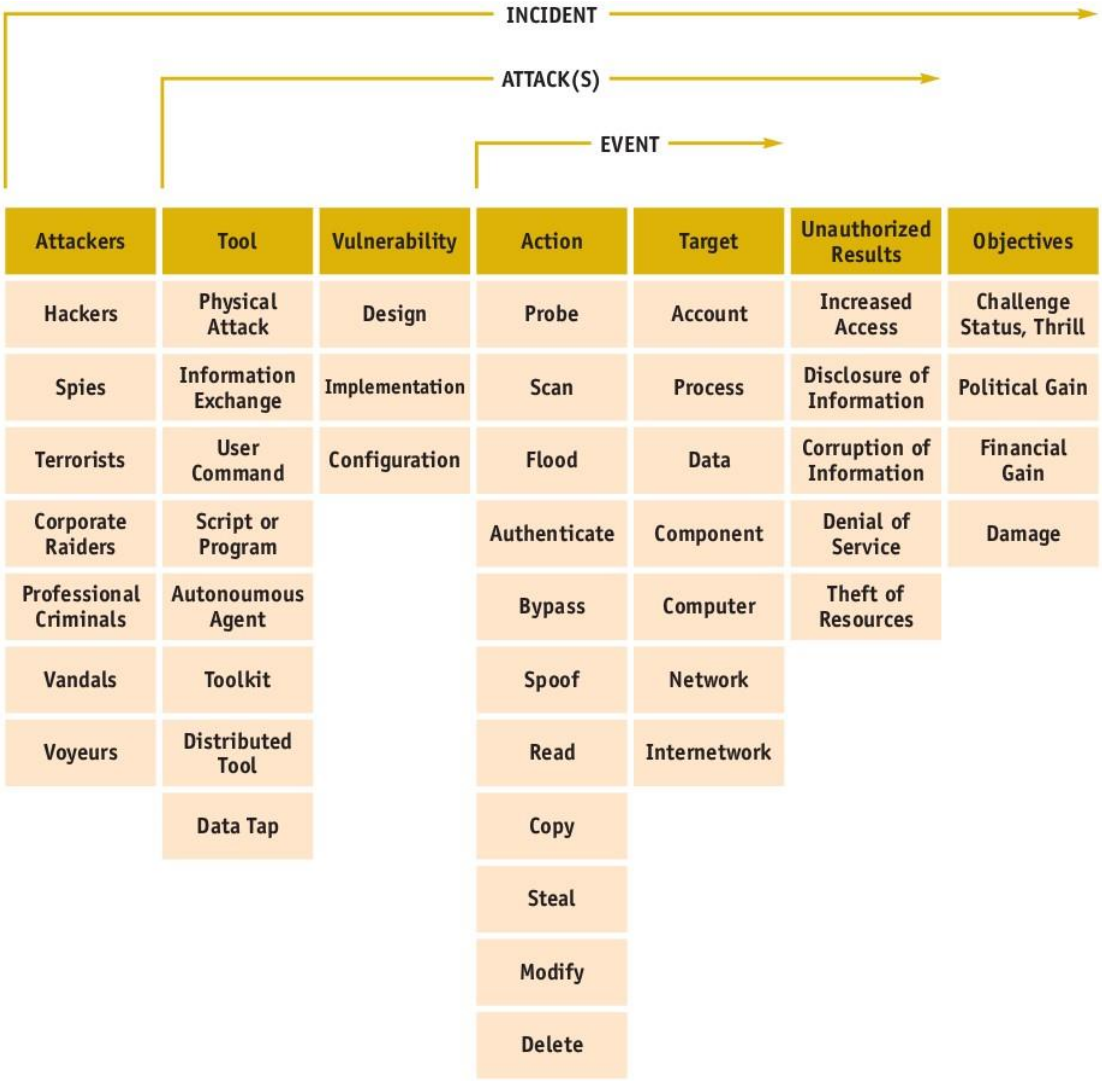


Figure 3. Common Language security incident taxonomy. Source: [15].

Besides verifying incidents, results of initial categorization and correlation should also provide information for initial classification of confirmed incidents (e.g. detection of spam, reconnaissance activity, or malware attack), and assign it an initial priority value (e.g. low, high, or critical). This is often performed in accordance with predefined schemes and criteria in use by a particular CSIRT [12], [16]. These schemes and criteria may vary depending on CSIRT mandate and its relations with constituency (e.g. public sector, private sector, or SLA). Initial classification and prioritization is important in terms of enabling further incident handling processes, since it identifies which CSIRT resources should be dedicated to handle the incident [15]. To accurately classify events as incidents CSIRT is required to answer several additional questions, such as [11], [15], [17]:

- Who has attacked us and what was attacker's motivation?
- Which tools and methods have been used to conduct the attack?
- What is the scope and the extent of the attack?
- When did the attack occur?
- Has any valuable information been compromised (see figure 3)?

Commonly a person performing triage does not have enough information to properly classify an event during this stage of incident lifecycle [12], [15]. For this reason incidents are often re-classified after detailed analysis is performed either before, or during the response process. Nevertheless, initial classification of an event is still necessary to enable further investigation. Any available information gathered during initial analysis helps to prioritize events and plan subsequent investigation activities more carefully [11], [15].

Implementation of triage functionality depends on several factors and may have certain dissimilarities among different organizations. Differences can be conditioned by the role that triage is expected to play in incident management process, staff that is assigned a responsibility to perform triage, skills and expertise capabilities available to a particular organization [12].

Triage can be assigned to a person, who is not required to have skills and knowledge that would allow him to estimate severity of incidents. Such low-level, or "tactical sorting" approach can be used, when triage is meant to be limited to categorization of incoming information, and other triage processes (incident verification, correlation and classification) are performed during incident response [12].

On the other hand, if triage is expected to provide precise assessment, or a high-level "strategic analysis" of verified incidents staff performing triage should have a combination of technical skills and understanding of risks related to specific business processes. Even though such approach requires an organization to devote a lot of support and training of triage staff, this allows to estimate the true impact of incident to organization during triage process, decreasing the time required to respond to incidents [12].

Triage can be also performed by organization's staff outside of a CSIRT (e.g. security officer, or IT help desk personnel). Using such triage implementation methods requires clear definitions of types, formats and means of information delivery to avoid a delayed response, caused by problems of communication between a CSIRT and staff performing triage. Delays in incident response can increase impact of incidents and the amount of damage they cause, or complicate further investigation process [12].

Regardless of triage implementation method, its effectiveness significantly depends on the quality of the information received from "detect" process. Lack of details required for adequate categorization and correlation of data can reduce accuracy and speed of triage, because CSIRT might be compelled to request additional information from the source [12], [16]. Next sections describe several measures that can be implemented by CSIRTs to enhance triage and avoid delays in triage process through improving quality of sub processes and processed information itself.

2.2.1 Incident verification.

Cyber-attacks commonly trigger multiple events on different security devices and software. These events may represent attack *precursors* (specific alerts, allowing to assume that an incident may occur in the future) and *indicators* (pieces of security-related information describing occurred or ongoing incidents). NIST has published a list of common sources of precursors and indicators together with alert descriptions in [11].

Detection and initial analysis of security events can be very difficult for several reasons. Since precursors and indicators are not guaranteed to represent a security event, ideally each of them should be analyzed to determine its accuracy. However, organization's security systems might record thousands or millions of possible signs of incidents on a daily basis, which makes their evaluation process extremely challenging. Even more confusion is created by the fact that sometimes accurate indicators do not necessarily mean that an incident has occurred, and sometimes indicators are rather hard to tie with potential incidents [11]. Advanced cyber-attacks can last for many months or years, often because attackers might have effectively covered their tracks. Even when such attacks are discovered, they are often assumed to be inappropriately classified [17]. Since determining if an event is actually an incident can be difficult, some organizations might establish their procedures, where events should be handled as incidents, regardless of

actual reason of an event [11]. Effectiveness of incident verification and its initial analysis can be increased by following several recommendations related to the "prepare" and "detect" processes, which are described in next sections.

System baseline analysis.

A CSIRT can profile networks and systems to determine changes in characteristics of expected behavior and support other detection and analysis techniques. Understanding normal behavior helps to recognize abnormal network and system activities [11]. Normal behavior of a host can be profiled through gathering information related to [18]:

- Open ports and processes. Applications and services allowed for usage within organization environment use certain TCP/UDP port numbers, which can be identified as acceptable.

- Running services and loaded drivers. Several default services and drivers should be started/loaded to allow functionality of host operating system and other required applications. Information related to acceptable services can be combined with the usage of ports and running processes to provide a better system setup baseline.

- User/Group information. Data gathered about legitimate user accounts and user group settings can help to identify unauthorized accounts, settings modifications, and thus prevent confidentiality compromise.

- Event logs and registry entries. Maintaining event log information provides a baseline for describing user interactions with various system elements in normally configured environment (e.g. by providing logon information, or system-specific errors). Data obtained from various logs can be confronted with security policies to detect possible incidents. Windows registry snapshots and backups can be used to determine unacceptable changes in operating system (e.g. installation of malicious programs and creation of processes).

Maintaining historical data.

Successful validation of incidents assumes that CSIRT performs event correlation among multiple indicator sources. Correlation of currently detected security events with data,

which has been historically captured by monitoring network and host activity, can assist in determining interconnections of different security events, and potentially detect long-lasting attacks. However, such correlation assumes that organization implements a log retention capabilities that would provide CSIRT with possibility to store historically recorded data for further usage. To improve correlation quality and avoid correlating inconsistent information about various events, it is important to synchronize clocks on hosts (e.g. by making use of maintaining NTP servers), predefine monitoring criterias, and filter out insignificant data [11].

Initial analysis of security events can be enhanced, if CSIRTs maintain a knowledge base of information, which allows to share explanations of specific precursors and indicators among CSIRT members. Such information repositories can be implemented in different forms, e.g. a website used for collaborative information sharing (wiki), or an incident tracking system, which is described in the following section. Usage of information sharing websites, or incident tracking systems creates possibilities for systematic collaboration with other resources (both organization internal and external) to obtain sufficient information related to a particular event [11].

2.2.2 Incident correlation.

Incident verification is a process of correlating various precursors and indicators with the purpose of determining attacker tools, methods and targeted systems [15]. As described in previous section, verified incidents can be additionally correlated with ongoing or past incidents to leverage incident analysis. Such correlation can identify organization security problems through determining deficiencies in organization security policies, infrastructure and employees awareness [12]. Correlation of incidents can be facilitated by maintaining knowledge base of historically gathered incident-related information.

Incident tracking systems is an example of information repository, which can support triage by storing information related to historically analyzed incidents, thus providing means for incident-related data correlation. Such systems complements incident related information with unique tracking numbers, which provide incident identifiers suitable for both, human and tool recognition. Besides using incident tracking numbers in a tracking system itself, they can be supplied during exchange of incident related information between triage processes, incident handlers, or organizations (e.g. by identifying specific

information in the subject line of email messages) [16]. Incident tracking systems can be configured to index specific fields of information contained in an incident ticket and add specific tags to improve data correlation capabilities.

Correlation of incident-related data can occur by searching tracking system for specific indicators, such as IP addresses or domain names, which can be stored in accordingly indexed data fields. Tags containing certain keywords can be applied to identify incident associations with a particular feature (e.g. adversary, targeted system, etc.). During "respond" process, a ticket related to a particular incidents can be updated with information about adversary TTPs, as well as their possible objectives and effective incident mitigation methods [19]. If such information stored as a result of previous investigations can be correlated, it can facilitate triage process through determining priority values to currently analyzed incidents.

2.2.3 Incident classification.

To support initial classification of incidents with regard to available information, classification schemas called "incident taxonomies" have been developed for individual CSIRTs and universal usage [15]. Researches of possible incident classification approaches have been conducted repeatedly. They have been mostly inspired by technological innovations, which created challenges in understanding the nature of cyber-attacks. Sophistication degree of cyber-attacks increased continuously, and every incident taxonomy related research addressed a specific problem. For this reason, different taxonomies have been developed to classify events, based on specific factors, e.g. [20] - [22]. Implementing a suitable taxonomy for classifying events in a CSIRT includes benefits and shortcomings, therefore is considered a recommendation, not a requirement [15]. Nevertheless, a research conducted by ENISA in 2016 determined that using a taxonomy provides CSIRTs with significant capabilities [23].

Tracking important metrics in CSIRT's constituency [23]. Incidents can be evaluated and triaged in terms of criticality of affected resources, the impact of an incident, the time and the effectiveness of processes used to solve them. Keeping historical information about resources required to solve incidents of particular types allows more adequate assignment of resources to current incidents. Even though the taxonomy as such doesn't necessarily affect evaluating incidents, it can support estimating incident severity by

providing tagging possibility. Tags can provide critical and actionable information by describing [23]:

- affected systems or victim type (e.g. governmental sector, or "critical infrastructure")
- law enforcement applicability (if relevant)
- sensitivity of incident-related information (e.g. is incident-related information allowed to be shared publicly, or should be processed by individual recipients?)
- Access restriction in cases of classified information (e.g. EU SECRET)
- Metrics for performance measurement (e.g. incident opening and closing timestamps, or incident status)
- Incident-related additional information (e.g. malware delivery method).

Facilitate initial classification of incidents [23]. Initial classification usually pre-defines allocation of organizational resources to handle incidents and assists in their prioritizing. If taxonomy consists of self-explanatory or clearly defined terms, the initial classification of incidents becomes simple and quick [23].

Exchange incident information with other CSIRTs [23]. CSIRTs may lack resources to perform exhaustive analysis of incidents and their effects on the system. In such cases the CSIRT may benefit from participating in cyber security information sharing network to cooperate with other teams, who have advanced technical capabilities in terms of incident analysis [11]. However, this assumes that participants of information sharing network use data standards and taxonomies that allow for synchronizing the processed information.

Support vulnerability management [23]. During incident management CSIRTs often have to process vulnerability information, such as system misconfiguration or something exploitable. Inclusion of corresponding category into a taxonomy would allow to share, report and resolve vulnerability issues more effectively.

Automate processing of incident information [23]. Clear definitions of terms within a taxonomy supplied with additional documentation allow incident-related data to be both

human-, and machine-readable. On the other hand, if categories within a taxonomy are not mutually exclusive, ambiguity of terms makes the machine reading harder.

2.2.4 Incident prioritization.

After the incident has been verified, correlated with other events/incidents and categorized, it has to be assigned a priority value. Since every organization is unique in terms of its business features, its information infrastructure can include very specific systems (e.g. weapon, industrial, or telecommunication systems) [24]. Thus, it is not uncommon that their incident prioritization schemas would have certain differences. However, for an effective incident management it is important to initially evaluate the incident after verification, independently of organization's peculiarity.

2.2.4.1 Factors that define incident severity.

Cyberattacks and different classes of cyber incidents are associated with specific risks that the organization should evaluate in terms of incident's influence of organizational operations, assets, individuals, other organizations, and the nation [24]. These issues need to be addressed from both the risk assessment and incident response planning perspectives. Risks implied by cyber incidents include [13], [25]:

- Physical safety risk: a cyber-attack against organization infrastructure could cause physical harm to individuals
- Reputational risk: cyber incidents may cause a negative impact on confidence in a system or provided service, public relations or cause legal proceedings with customers
- Regulatory risk: cyber incidents may result in violation of a regulations related to data processing requirements, established by relevant legislation
- Operational risk: cyber incidents can cause disruptions to critical business operations
- Financial risk: cyber incidents can cause financial losses associated with lost data, stock price dropping, and the loss of physical assets
- Internal human relations issues related to payroll and employee privacy.

Evaluating severity of a particular incident can be a challenging task, where multiple factors have to be taken into consideration. As a possible solution to such complicated problem ITIL proposes to consider two main factors during incident evaluation [26]:

- incident's impact on users and on the business
- incident's urgency,

where *impact* is defined as "the measure of the extent of potential damage the incident may cause" [26];

while *urgency* indicates "how quickly a resolution is required" [26].

A distinctive feature of ITIL is that it describes IT-related procedures in context of service availability from the perspective of system users and customers. Hence, besides the aspect of data availability, the fundamental concepts of information security also include confidentiality and integrity of data, which are not focused by ITIL. This means, that ITIL method of incident evaluation can't be considered exhaustive, at least in context of responding to the cyber threats. Thus, evaluating significance of security incidents requires another methodology.

Another approach is proposed in a research conducted by the NIST [11]. According to NIST recommendations, effective incident handling assumes that security incident should be prioritized considering 3 following factors:

Functional impact of the incident, which reflects capabilities of organization to provide services to users and customers with regard to possible incident escalation (see table 1).

Table 1. Categories of incident functional impact. Source: [11].

Category	Definition
None	No effect to the organization's ability to provide all services to all users
Low	Minimal effect; the organization can still provide all critical services to all users but has lost efficiency
Medium	Organization has lost the ability to provide a critical service to a subset of system users
High	Organization is no longer able to provide some critical services to any users

Information impact of the incident, which represents risks related to unauthorized access to sensitive information and its possible exfiltration (see table 2).

Table 2. Possible categories of incident information impact. Source: [11].

Category	Definition
None	No information was exfiltrated, changed, deleted, or otherwise compromised
Privacy Breach	Sensitive personally identifiable information (PII) of taxpayers, employees, beneficiaries, etc. was accessed or exfiltrated
Proprietary Breach	Unclassified proprietary information, such as protected critical infrastructure information (PCII), was accessed or exfiltrated
Integrity Loss	Sensitive or proprietary information was changed or deleted

"Recoverability from the incident, determining the amount of time and resources that must be spent on recovering from the incident" (see table 3).

Table 3. Examples of recoverability effort categories. Source: [11].

Category	Definition
Regular	Time to recovery is predictable with existing resources
Supplemented	Time to recovery is predictable with additional resources
Extended	Time to recovery is unpredictable; additional resources and outside help are needed
Not Recoverable	Recovery from the incident is not possible (e.g., sensitive data exfiltrated and posted publicly); launch investigation

Such model of incident evaluation is developed to assist in assigning priorities based on business impact (which is represented as a combination of functional and informational impacts), and cost of recoverability (an aggregate of time, human and financial resources). The incident with a greater impact on business and lower recovery costs should be assigned a higher severity value over incidents with a smaller business impact and higher recovery costs [11]. Although business companies and governmental structures might have unique security policies, such general approach is designed to be universally adopted. Presuming enough flexibility to meet organizational specificity, it allows to describe impacts of a particular incident. On the other hand, this approach represents a high-level (strategical) terminology, but doesn't provide methods of evaluating functional, information impacts and recoverability from the incident. For this reason, a

lower-level (tactical) methods of estimating business impact of incidents are described in the following sections.

2.2.4.2 Threat severity estimation.

Triage of incidents can be facilitated by supplying CSIRT with information about organization relevant threats, internal and external vulnerabilities, and potential impact to organizations that may occur if vulnerabilities are exploited by threats. In context of incident management, such analysis, called cyber risk assessment, should be continuously performed throughout the system development life cycle [24].

"Threat event" can be defined by organizations as single security events, malicious actions, or circumstances; or sets of related events, actions and circumstances with the potentially adversely impact. Threat event examples can include [24]:

- Adversarial threat events:
    - Perimeter network reconnaissance or scanning
    - Receiving phishing email messages
    - Malware delivery to internal information systems
- Non-adversarial threat events:
    - Natural disasters
    - Mishandling of critical and/or sensitive information by authorized users
    - Disclosure of software vulnerabilities

Such threat events are caused by threat sources, which can be characterized as the intent and method, used to exploit system vulnerability; or the situation and method that may accidentally exploit system vulnerability. Types of threat sources can be generally divided into [24]:

- Malicious activities performed by adversaries (e.g. competitors, customers, or nation-states), including:

- Outsiders
- (Trusted or privileged) insiders

- Legitimate system users omission and commission errors
- Hardware or software failures
- Disasters, including both natural and man-made

Adversarial threat events can be characterized by the TTPs used by the adversaries. Understanding TTPs provides a better understanding of possible attack objectives and helps to focus on a specific set of threat events that are most relevant to organization [24]. In context of determining incident severity, threats that are relevant to the organization should be measured to enable quantification of the corresponding incidents.

Author of this research used publicly available contact list of CSIRTs with national responsibility[1] to conduct a questioning about models that are in use for estimating threat severity. Although only a few CSIRTs responded to author's questions, majority of responses stated that threat severity is commonly determined based on a common sense and experience obtained through resolving incidents. However, SANS has described a more systematic approach, proposing to conduct threat measurement in terms of threat source motivation and capabilities [27]. Regardless of a selected model, determining severity of a particular threat would assist in estimating incident severity, thus facilitating triage process.

2.2.4.3 System vulnerabilities.

A vulnerability is a weakness in an information system itself, security procedures, internal controls, or implementation methods that could be exploited by a threat source. In regard to cyber domain, such security weaknesses may result from design of a particular product

[1] *CSIRTs with National Responsibility Contact List* [Online]. Available: https://www.academia.edu/5913580/CSIRTs_with_National_Responsibility_Contact_List?auto=download

(errors in application source code), security controls that either have not been applied (e.g. security patches), or have been applied, but still contain some weaknesses (e.g. system misconfiguration implemented by administrators) [24], [28]. However, vulnerability identification is not only limited to cyber domain, they can be also determined in [24]:

- organizational management (e.g., insufficient knowledge of critical communication channels)
- external relationships (e.g., low diversity of risks in regard to energy, or technology dependencies)
- business processes (e.g., poorly defined processes and responsibilities)

Since vulnerabilities create possibilities to exploit systems and get unauthorized access to sensitive information, they should be considered as severe security risks [29]. Severity of information system vulnerabilities can be described with the help of CVSS. CVSS can support triage of incidents by providing a numerical score that summarizes different metric groups. Description of these metrics is out of the scope of this research; however, important feature of CVSS is that it includes metrics, which describe confidentiality, integrity and availability impacts of a successfully exploited vulnerability [30]. These values are specified as “High” (“H”), “Low” (“L”), or “None” (“N”) and can be compared to the corresponding values, associated with a particular host protection requirements (described in further section) to estimate impact of a specific threat [31].

During triage of security events, information system vulnerabilities should be evaluated in combination with predisposing conditions. A predisposing condition is a condition which affects (i.e., increases or decreases) the likelihood that a threat can exploit a vulnerability to cause negative impact on organization. Some examples of predisposing conditions include [24]:

- Geographical location (e.g., datacenter, located in highly-seismic region is more likely to be affected by earthquakes)
- Specific characteristics of information systems (e.g., networks isolated from external connections are less likely to be exposed to network-based attacks, while

absence of antivirus software increases successfulness of malware-based attacks attempts)

Regular vulnerability scans performed within organization information system environment can provide CSIRT with relevant information about every particular host. Maintaining such information allows to estimate the likelihood of a specific threat to cause an incident, and if incident occurs - estimate its functional and information impacts in context of confidentiality, integrity, or availability breaches.

2.2.4.4 Incident impact estimation.

Once a cyber incident is determined it can be thought of as a risk to an organization with a potential chance of its realization. The likelihood of risk realization depends on a combination of threat and vulnerability characteristics, predisposing conditions affecting the likelihood of threat to exploit the vulnerability and effectiveness of implemented security controls in terms of mitigating a particular threat (see figure 7).

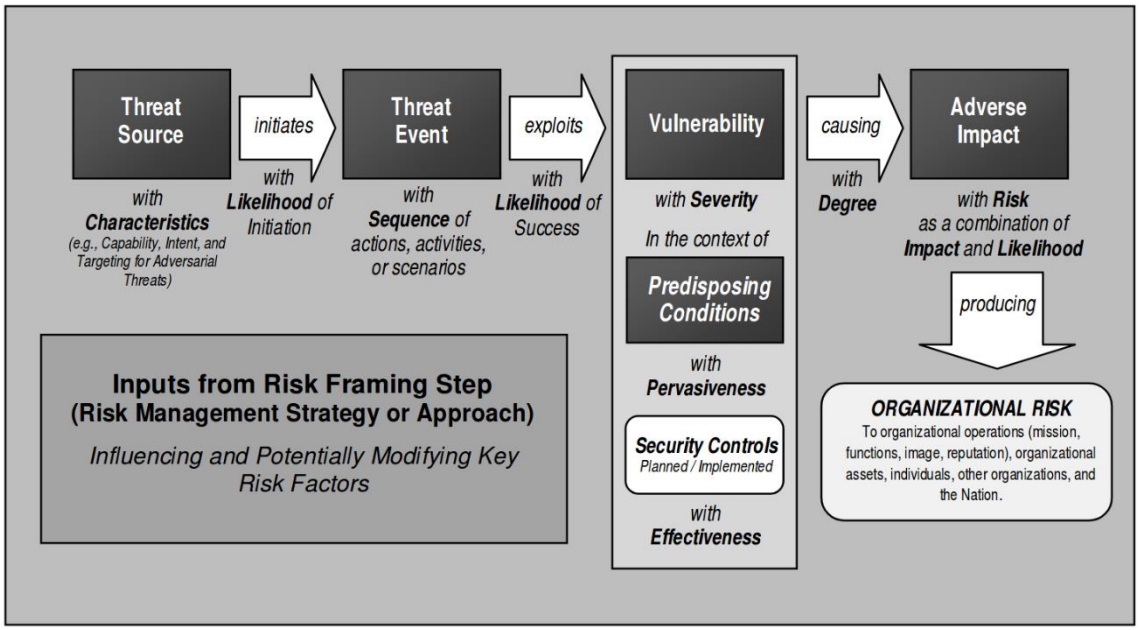


Figure 4. Relations of factors affecting risk assessment. Source: [24].

As long as predisposing conditions and implemented security controls successfully prevent threats from exploiting vulnerabilities (e.g. firewall drops unwanted ingress traffic), severity of determined incidents can be considered as “low”, or “good to know”. But whenever security controls fail to protect the targeted systems, incident severity increases and estimating the impact becomes important. Impact of incidents can be

estimated through collaborative analysis of the previously described factors and additional adjustment, which depends on the characteristics of affected assets.

Functional impact. As described in previous sections, certain threats might be associated with specific severity values that can be determined either by risk assessment, or experience historically obtained by CSIRT. Such values can be pre-defined as integers, representing relevance and significance of malicious activity (e.g. "1" representing a low severity threat, and "10" for critical threats affecting systems).

However, the functional impact of a particular threat should be adjusted considering affected system's criticality in terms of supporting business functions, its location within network and its interconnections with other hosts. Regardless of organization business functions (e.g. industrial, commercial, or military sector), certain assets typically have a more significant importance than others. CSIRT can't expect that the same threat event would affect different hosts in the same manner. This means that threat event severity should be evaluated in accordance with the information about affected hosts [31]. For example, if organization relies on its internal DNS servers, prevention of its availability would affect all other services that rely on name resolution (e.g. email, web-based services, etc.). On the other hand, a compromised workstation of a non-administrative employee might prevent him or her from performing daily duties, but such incident is not likely to cause additional impact to other users and services (if adequate security mechanisms are in place).

Thus, a "host importance" factor should be considered for the purpose of adjusting threat severity values in context of estimating functional impact of incidents. This factor can be calculated through evaluating host *exposure to vulnerabilities* and *criticality of host-related assets* (e.g. installed applications, or hosted data) [31]. Host's exposure to vulnerability can be estimated with the help on a particular vulnerability CVSS score, or a function of CVSS scores, if host contains multiple vulnerabilities. Estimating asset criticality can be a challenging task because of uniqueness of organization business functions. However, instead of considering specific characteristics of every unique host, its criticality can be associated with the importance of network segment, that host is belonging to. Such approach would allow to estimate host importance, represented as an integer value, without a need to manually collect data about specific host-related attributes.

Organization's cyber environment typically consists of several network segments with different criticality measure in context of supporting business functions and protecting information. Some examples of network segments may include [32], [33]:

- DMZ. Such systems are designed to be exposed to the Internet. Typical examples include organization's websites, or mail gateways [32], [33];
- Business network. This segment houses end-user workstations, servers, printers and other systems supporting non-core management [32], [33].
- Business network management. It contains administrative user workstations, configuration and log management servers, as well as security management systems [32], [33].
- Critical system zone. The critical system zone houses systems that operate critical processes and database servers. A breach of confidentiality, integrity, or availability of data processed by such systems, or hosted in such databases is likely to be associated with realization of multiple types of risks described previously [32].
- Safety systems segment. Malfunction of safety systems can cause damage to operational environment, or injuries to personnel. Examples of critical safety systems include: fire suppression, electricity generation, or nuclear reactor control systems [32].

Organization may develop and implement a unique scoring system that can use integers to reflect the criticality level of incidents determined in particular network segments. Affected hosts can be easily associated with a particular network segment with the help of assigned IP addresses. Such approach would allow to evaluate incidents according to organization's unique requirements. For example, a particular security threat determined in DMZ can be assigned a score of "1", while the same threat affecting database containing personal information of customers (located within critical system zone) could be evaluated as "10".

Prevention of access to a particular host and services that rely on its functionality in context of information security can be thought of as information availability breach.

Alternatively to estimating criticality of assets in accordance with criticality of network segments that house the data, each host can be assigned a more accurate priority value, representing its significance in terms of providing access to relevant information. In addition to determining host criticality more accurately, such value (e.g. "1" for regular user workstation, or "10" for internal DNS server) can be compared against corresponding vulnerability metrics to incident availability impact more specifically, if a vulnerability would be successfully exploited [31].

Information impact. An approach of estimating incident information impact proposed by NIST describes categories of risk, related to unauthorized access to sensitive information. However, it is important to mention that proposed categories are not mutually exclusive and during estimating incident severity should be assessed independently [11]. A disadvantage of this model is that it doesn't take into account different classes of information, which is a significant factor affecting incident severity in military, or governmental organizations (e.g. unauthorized access to "NATO CONFIDENTIAL", U.S. "Top Secret", or "EU RESTRICTED" data).

Information impact can be estimated using similar methods used for evaluating functional impact: either through determining host location in a particular network segment, or individually, if necessary. However, instead of calculating data availability factor, information impact should be estimated with data confidentiality and integrity criticality values. Although modification or exfiltration of affected information can be confirmed during incident response, triage should determine potential impact of information confidentiality, or integrity breaches. Several standards and recommendations [34]-[36] can assist organization in estimating values of different data types, to support evaluation of related incident severity.

Severity of data confidentiality, or integrity breaches can be represented as a function of corresponding asset criticality value and vulnerability metrics. Given that confidentiality and integrity related numerical values, that characterize severity of incident are not likely to be the same (because of differences in asset criticality and vulnerability metrics), it should be decided, which value to use for triage purposes. One possible method is proposed by the "Three-level IT Baseline Security System "ISKE"" [37]. It suggests that value, describing information security requirements should be determined by the highest value among confidentiality, integrity, and availability metrics associated with host

importance. The benefit of such method is ease of its implementation, however, the accuracy of result wouldn't necessarily correspond to organization requirements.

To improve accuracy of estimating information impact, another research [31] proposes to calculate it as the Euclidean norm of confidentiality and integrity values. However, this approach doesn't consider functional impact as an individual factor, and for this reason may be unacceptable for certain organizations that rely on information systems with high availability requirements.

3 Threat intelligence and its role in hunting cyber threats.

3.1 Threat intelligence sharing aspects.

TI is a relatively new term in cyber security field, and for this reason it has been used to describe different products, services, or even processes [38]. However, in general the term describes various kinds of information related to cyber threats, which can be used to support decision-makers upon its disclosure. The aim of processing TI is to detect cyber incidents in a timelier manner, or event prevent them from occurring [39]. Several researches conducted by SANS determined that popularity of processing TI by security professionals has shown a significant growth beginning form 2014. In 2016 only 6% of questioned respondents confirmed that they haven't used TI for securing their cyber environment [40].

The demand for exchanging TI with other organizations increased in recent years, since it provides analysts with access to information that might otherwise be unavailable. Cooperation in the field of TI sharing enables organizations to enrich the collective knowledge and analytic capabilities allowing "one organization's detection to become another's prevention". Participating in TI sharing communities offers companies many advantages, which can be summarized into the following definitions.

Shared Situational Awareness. Even a single contribution can increase can the awareness and security of an entire cyber security community [41].

Improved Security Posture. Using shared information organizations can identify potential targets within their environment, implement proactive protective measures, improve intrusion detection rates, and respond to the incidents more effectively [41].

Knowledge Maturation. Observations that initially appear insignificant can be correlated with data collected by other analysts. Such correlation can enrich existing information, develop knowledge base related to a particular threat, and determine relationships between different indicators, associated with a specific threat campaign [41].

Greater Defensive Agility. Organizations that share TI are better informed about recent TTPs used by adversaries, which helps to reduce probability of successful attacks [41].

Even though aforementioned benefits are obvious, organizations that are willing to develop a TI program should consider several challenges, related to processing shared intelligence. These include legal, organizational, and technical nuances.

Legal challenges. Insufficient safeguards may result in unauthorized disclosure of sensitive and classified information and lead to financial loss, violation of sharing agreements, legal proceedings and loss of reputation [41]. Permissions to process classified information are issued only to specific organizations, which meet established criteria and moreover - on a per-person basis. Thus, getting such permissions can be expensive and time-consuming process. Law enforcement is an example of governmental organization, which frequently release TI, however such information released by law enforcement is not always meant for public distribution [42].

Organizational nuances. A significant effort is required to establish and maintain trust relationships that form the foundation for any sharing processes. This process can be enhanced through regular cooperation (e.g. meetings, or phone calls) with sharing peers. In addition, creating data sharing capabilities assumes that organization has the necessary infrastructure, tools, personnel and training available. Insufficient training, or irresponsible sharing of sensitive information can expose data related to organization's internal security solutions and lead to threat shifting, disruption of investigation or response actions [41]. It is also important to evaluate effectiveness of potential security feeds (commercial vs free), since TI obtained via commercial feeds are not necessarily the most useful [38]. Moreover, data sources should be chosen carefully, since investments into purchasing threat intelligence can be rather expensive, but not always cost-effective [38], [42].

Technical considerations. Even if CSIRT is provided with the required resources, it has to carefully estimate its processing capabilities to avoid receiving overwhelming volumes of data [41]. This becomes important, because the deployment of advanced security sensors and defenses have caused a significant increase in volumes of data processed by TI tools [43]. Prior to reacting to a received TI, an organization needs to evaluate its quality, relevance, and understand the risks of using or not using the information [41], [45]. Quality of TI can be determined by reputation of information source, and latency between a particular threat detection and release of corresponding TI. External sources of TI should be validated (e.g. by using hashes and digital signing) to ensure that data

originates from legitimate source and has not been modified in transit [43]. Avoiding delays between information sharing partners can be achieved by using standardized data formats and transport protocols. Adopting such formats can require significant resources, but benefits of investing into such resources can be considerably reduced, if partners operate with different data standards or protocols [41], [43], [44].

If an organization decides to participate in exchanging TI, the goals and objectives should be evaluated together with the available resources and described challenges to determine the desired outcomes of this process [41].

3.2 Types and sources of threat intelligence.

Chapter 2 described cyber incident management as a collaborative workflow that assumes multi-layered cooperation of organization's personnel. Considering generality of a given definition of threat intelligence, it may include vast amount of multifarious information, including analysis of policy releases, cyber security white papers, technical information, etc. This means that the same principle should apply to processing TI, since effectiveness of its analysis assumes that specific information should be processed by decision makers with corresponding competence. For example, high-level managers (CEO, CFO, CIO, etc.) should be aware of possible cyber risk displacements that may occur in result of trend and observation assessment. On the other hand, they are not likely to deal with particular email message subjects or file hashes, though such types of TI can be immediately applied by technical personnel for threat hunting purposes [38].

Research of TI area conducted in 2015 proposed a framework for identifying subtypes of TI to support effective implementation of TI programs [38]. According to the result of research, consumed intelligence can be divided into subtypes in accordance with its content (impacts on business related decisions) and corresponding consumers of data. Proposed subtypes include strategic, operational, tactical and technical TI.

Strategic TI. As mentioned previously, strategic TI is important for organization high-level strategist (board level directors and chief level officers) to estimate current cyber risks and identify possible shift of risk in future [38]. Understanding actual threats to the business allows to allocate organizational resources and implement new technologies to adequately protect critical assets and business processes [45]. Strategic TI can be

produced through analysis of information with strategic importance, which can be obtained from various high-level information sources including [38]:

- Geopolitical assessment (e.g. analysis of national policy releases, or news in subject-specific press)
- Security industry white papers that may contain relevant information about cyber threat landscape
- Trustworthy human peers in relevant organizations that can provide valuable information about ongoing cyber campaigns and current threats.

Strategic analysis is a challenging process, which requires good understanding of the sociological and political aspects. Intelligence produced through strategic analysis is often represented in form of conversations, briefings, or reports, developed for long-term usage. Strategic TI is not usually shared because of its sensitive nature [38].

Operational TI. Another high-level type of TI is operational intelligence, which provides actionable information on specific cyber-attacks and adversary capabilities. Such information is useful for higher-level security managers and leaders of incident response teams in terms of providing situational awareness [38] and context for tactical TI [43]. Operational TI may define probable actions of adversaries in regards to organizational infrastructure, and help to estimate effectiveness of investments in security products [38], [45]. However, collecting operational TI is usually problematic for private companies because of insufficient access permission to relevant data channels and repositories, which can be used by the governmental structures. Alternatively, private organizations might gather operational information from [38]:

- Correlation of recent cyber-attacks with real-world events. Analysis of recent attacks, related to specific adversary groups may help to determine interconnections between real world indicators (e.g. specific social media posts) and subsequent cyber activities.
- Open communication channels. Some adversaries use chat rooms with unrestricted access to discuss and coordinate their activities. However, adversaries that are aware about risks of being monitored in such chat rooms, use private

rooms to prevent information disclosure. Some vendors sell TI obtained from such private chat rooms; however, buyers would need to ensure that this product is obtained legally.

- Social media. It is possible to gather operational information through monitoring social network for matches of specific keywords (e.g. name of specific company). Some vendors can produce operational information as a streaming feed of posts, which contain such predetermined matches.

The best way to produce operational TI is to implement automated solutions for analysis of activity- or event correlation results. This is rather challenging because gathered information can be obtained in different languages, in the form of slang, or aliases used by adversaries. Operational TI can be shared with other organizations and therefore can be also obtained from others [38]. Private sector organizations working in the same industry segment (e.g. financial services, or information technology) might exchange information about security threats with the support of ISAC groups [42]. ISACs are nonprofit organizations, which have been founded worldwide to enhance cyber security awareness by sharing security-related information between their members.

Tactical TI. While high-level TI aims to estimate business risks and predict whether a particular organization is likely to become a target of a specific type of cyber-attack, tactical TI can be used to describe adversary TTPs. Tactical TI provides actionable information for detecting and mitigating attacks, thus, it is a valuable data for system architects, administrators and security personnel. TI is commonly gathered either by security sensors [43], or in a form of various reports, including [38]:

- Attack group or campaign reports, which are most common sources of tactical TI (e.g. [46 - 48])
- Analyzed malware samples and analysis reports released by security community
- Incident reports that can be published formally, or obtained from peer-defenders or -investigators in informal manner.

During processing of tactical TI analysts should extract patterns of adversary behavior, identify system vulnerabilities that are exploited by adversaries and determine, whether

such vulnerabilities are present in organization's environment [38]. Exchanging tactical TI is strongly recommended, since it encourages the community to contribute into knowledge maturation, improve community situational awareness, security posture and defense agility [38], [41].

Technical TI. Technical TI is another low-level type of actionable information, which can be described as specific indicators that can be quickly distributed and included in defensive infrastructure. Technical TI is often referred to (but not limited to) as IOCs and it is consumed by CSIRTs to detect or prevent cyber incidents from occurring [38]. With the help of technical TI incident management can be facilitated by automating event correlation, decreasing false-positive alarm rates and supporting prioritization of vulnerability patches and security events [45]. A detailed description of IOC examples, collection methods and sharing methodologies is provided in the following section.

Although TI can be divided into different subtypes for more efficient processing (see figure 5), it is important to understand that in result of analysis performed on one of the levels, specific elements of intelligence can be shared with other levels to support workflows of different analysts. For example, tactical TI can be used to extract indicators, which are most valuable for usage on technical level. Similarly, strategic analysis may produce actionable guidelines for each lower level in regard to cases when specific types of malicious activity is detected.

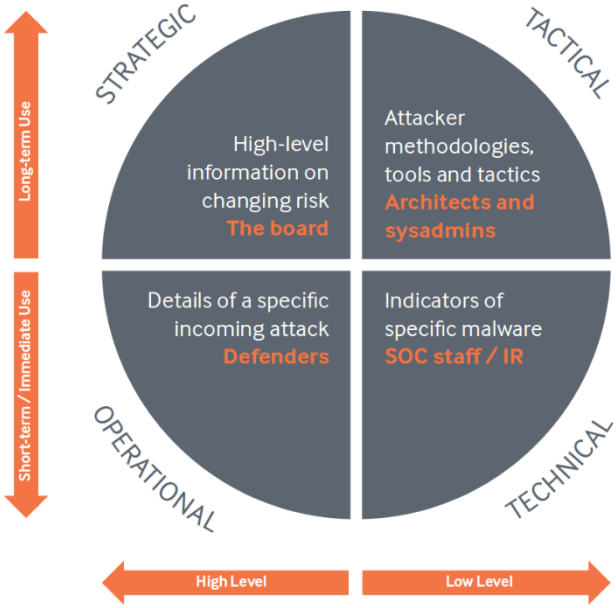


Figure 5. Subtypes of threat intelligence. Source:[38].

3.3 Indicators of compromise and IOC sharing methodologies.

3.3.1 IOCs and their quality.

The structure of cyber-attacks can be divided into several stages, which form up a process commonly known as the "Cyber Kill Chain" [40], [49], [50]. Understanding kill chain and attacker's activities during different attack stages allows to ensure that if some defense measures are bypassed by the adversaries, other defense mechanisms would still determine the intrusion [40]. Some attack-related activities can be performed by the adversaries offline (e.g. passive system reconnaissance, or evaluating possible attack vectors and effective payloads), while certain actions assume interacting with a target system. Whether attackers actively scan a system, send phishing messages, or exchange data between a compromised system and their C&C servers, such actions leave digital traces in network and end-host devices [50]. Disclosure of such traces, related to malicious activities compels adversaries to change their tactics, tools or approach methodology in order to accomplish their objectives [51].

IOC is a piece of forensic data, which consist of digital footprints or artifacts, associated with a particular adversary TTPs [42]. IOCs can be thought of as technical descriptors of attacks and compromise that can be used by CSIRTs to conduct incident investigations and perform intrusion detection activities [52]. The term "indicator of compromise" was first used by government organizations and defense contractors, who were dealing with identification of advanced persistent threats. Cyber security industry professionals have started using the term widely since 2007 [51]. Examples of the most common types of artifacts, which can be used for developing IOCs include file hashes, IP addresses, domain names and registry entries.

Advantage of using IOCs is that it allows to assemble multiple artifacts with the help of simple and complex expressions for the purpose of determining, whether security of the network has been breached [49]. Some IOCs may be developed from a single artifact, once it has been confirmed to be associated with a specific malicious activity, or adversary group. Such IOCs can be classified as atomic (e.g. IP address, or domain names), or computed (file hashes, or IDS signatures) [53]. However, a qualitative IOC represents a combination of several digital artifacts, found in log entries, system files, or network traffic [49]. Such collection of indicators can be described as behavioral IOCs [53]. In

contrast with analyzing individual pieces of digital evidence, searching for behavioral IOCs facilitates CSIRTs by increasing threat detection and decreasing false positives rates. This is because behavioral IOCs is confirmed to be detected only when a set of multiple atomic and computed IOCs have been detected within an environment [42], [51].

A vivid graphical illustration that highlights disparity of IOC values has been published by David Bianco [54], who created a diagram called "The Pyramid of Pain" (see figure 5). From the defender perspective the diagram shows the relationships of different types of indicators, their effectiveness and complexity of development. On the other hand, the diagram represents relationships between indicators and adversary efforts required to modify their TTPs, if a particular indicator has been disclosed by defenders.

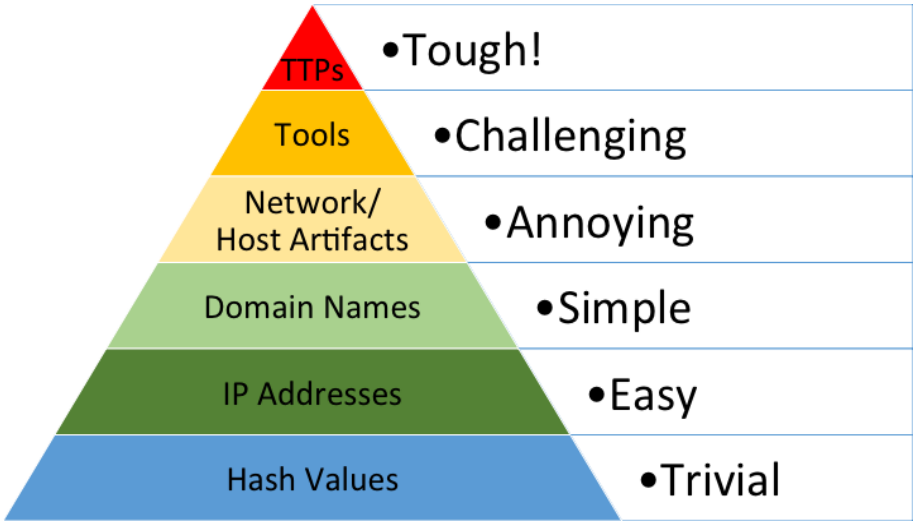


Figure 6. The Pyramid of Pain. Source: [54].

File hashes are easy to obtain from monitoring tools and calculate whenever required; however, they are also trivial to modify because of computing algorithms design (a single null-byte added to the end of a binary will result in a totally different hash value). For this reason file hashes can't be considered reliable IOCs, when used individually. The same applies to IP addresses, since advanced adversaries are very well aware of different anonymous proxy services, allowing them to change IP address whenever necessary. Migrating from one domain name to another can be insignificantly harder for adversaries, but given the fact that many DNS providers offer hosting services free of charge, this can't be considered a serious challenge for attackers [54].

The first level where the defenders can cause some perceptible impact on the adversaries is the "Network/Host Artifacts". Detection and reaction to disclosure of network-, and

host-based artifacts allows to compel adversaries to spend some time analyzing implemented countermeasures and adjusting their tools. On a higher level antivirus or Yara signatures can be used as tool indicators, allowing to describe functionality of the tools, instead of their static parameters. Implementation of such indicators can force adversaries to search for alternative tools, develop new ones, and spend time on figuring out how to use them effectively [54].

The TTPs are located at the top of the pyramid. At this level, the response is directed at adversary behaviors, not against their tools. If such response is implemented effectively, adversaries are forced to reinvent attack approaches, which is the most time-consuming process from their perspective [54]. TTPs of adversaries can be described using behavioral IOCs. While it is almost trivial to extract static and create computed IOCs during malware analysis, developing intelligence that would indicate to a specific adversary is challenging. It makes behavioral IOCs most valuable from defenders perspective; however, they are most difficult to obtain, or develop [52]. To increase effectiveness of behavioral IOCs the following guidelines can be considered during its development [51]:

- Specific areas of the operating systems that attackers commonly use during intrusion (e.g. file system, registry) should be examined for possible changes, even if exploitation of a particular system occurred using different intrusion paths.
- A developer should distinguish artifacts left by attacker's tools that would be inexpedient to change or modify for financial or other reasons.
- Systems that were not directly compromised can be searched for unusual activity, which could identify attacker's lateral movement techniques.
- In cases of critical systems with limited activities use whitelisting principles instead of blacklisting data. For example, an IOC developer can make a whitelist of legitimate files located in a specific directory and inspect all files not on that list.

Even behavioral IOCs are not identical in their values. Apart from estimating effectiveness of IOC in terms of its potential impact on adversaries, the quality of IOCs

can be determined with the help of true positive and false positive incident detection indices. Consider two following examples of behavioral IOCs:

1. Deploy a malicious file A to establish communication with IP address A.B.C.D.

2. Use server located at domain ABC, or XYZ to send an email message with title "123", and attached PDF file with embedded malware. Upon execution, drop files XYZ into directory ABC, with hidden directory attributes set. Execute file A, or B, or C to start outbound network connect to TCP port 53, or 80 at domain ABC, or XYZ. Download DLL file ABC and modify registry entries ABC and XYZ [52].

The described examples demonstrates how individual atomic (domain and directory names, port numbers, file extensions and registry entry values) and computed (file hashes) IOCs can be assembled to provide additional context to individual artifacts, describe the entire attack kill chain, and decrease false positive alert rates by providing more specific matches.

3.3.2 Formatting IOCs.

Effectiveness of high-quality behavioral IOCs can be significantly decreased, if defenders can use them only for cleanup processes, instead of preventing incidents from occurring [43]. Depending on the speed of technical TI distribution, changing nature of cyber threats can reduce its value to zero in days or even hours. In recent years cybersecurity community developed many standards and tools for storing and exchanging TI to improve its managing and sharing processes [55]. ENISA research showed that all of them have certain advantages and disadvantages (that are described further), however none of them can be considered as commonly accepted [56].

IODEF and IODEF-SCI. In December of 2007 R. Danyliw, J. Meijer and Y. Demchenko defined a standard for exchanging security information between CSIRTs[1].

[1] *The Incident Object Description Exchange Format, RFC 5070* [Online]. Available: https://www.ietf.org/rfc/rfc5070.txt

"The standard is implemented in XML language, allowing to encode information about hosts, networks, and the services running on these systems; attack methodology and associated forensic evidence; the impact of the activity; and limited approaches for documenting workflow". IODEF extension that supports description of additional data has been developed by T. Takahashi, K. Landfield, T. Millar and Y. Kadobayashi in July of 2013. IODEF for Structured Cyber security Information (IODEF-SCI) enables to embed and convey other structured cybersecurity information to enrich IODEF data and facilitate exchange of TI[1]. IODEF-SCI allows to include the following standards:

- Common Attack Pattern Enumeration and Classification (CAPEC)
- Common Configuration Enumeration (CCE)
- Common Configuration Scoring System (CCSS)
- Common Event Expression (CEE)
- Common Platform Enumeration (CPE)
- Common Vulnerability and Exposures (CVE)
- Common Vulnerability Reporting Format (CVRF)
- Common Vulnerability Scoring System (CVSS)
- Common Weakness Enumeration (CWE)
- Common Weakness Scoring System (CWSS)
- Malware Attribute Enumeration and Characterization (MAEC)
- Open Checklist Interactive Language (OCIL)

[1] *An Incident Object Description Exchange Format (IODEF) Extension for Structured Cybersecurity Information, RFC 7203* [Online]. Available: https://tools.ietf.org/html/rfc7203

- Open Vulnerability and Assessment Language (OVAL)
- Extensible Configuration Checklist Description Format (XCCDF)

The following benefits and shortcomings of IODEF have been determined during research conducted by ENISA [56]:

Benefits:

- IETF Open Standard defined by CERTs and for CERTs
- Enables a collaborative effort
- Vendor neutral in origin
- Flexible format (XML) allowing for extensions and the grouping of events data
- Allows for the grouping of events data

Shortcomings:

- Limited adoption
- Incident data can contain sensitive information harder to share
- High granularity that can complicate implementation

RID. While IODEF and IODEF-SCI describe standards for security data encoding, the Real-time Inter-network Defense (RID) provides a secure method to exchange information contained in IODEF documents[1]. RID integrates detection, tracing, source identification, and mitigation mechanisms to enable exchange of potentially sensitive information. Similarly to IODEF and IODEF-SCI, RID uses XML to encode its messages, which simplifies its integration with other aspects for incident handling.

[1] *Real-time Inter-network Defense (RID). RFC 6542* [Online]. Available: https://tools.ietf.org/html/rfc6545

Security of communication is achieved by utilizing TLS, XML security features of encryption, and digital signatures.

Advantages of RID include [56]:

- Developed, reviewed, and published by the IETF, which ensures its capabilities
- Existing open source implementations have successfully passed interoperability tests
- Provides decent level of information confidentiality, integrity and source authentication

Disadvantages of RID [56]:

- Utilized security mechanism are designed for peer-to-peer communications, which limits RID adoption
- High granularity that can complicate implementation
- Security options can lead to high implementation costs
- Not used in practice [57]

ROLIE. As an alternative to RID, in September of 2012 J. Field from the MILE Working Group introduced the Resource-Oriented Lightweight Indicator Exchange. ROLIE was developed as a more agile solution for exchanging TI broadly (as Web-addressable resources), instead of establishing peer-to-peer trust relations. The transport protocol binding for ROLIE is specified as HTTP(S) with a media type of Atom+XML[1]. Despite of effort placed into development of ROLIE, a study conducted by ENISA in November of 2014 determined that this standard is no longer maintained and hasn't found possibilities to be utilized by any tools [57].

[1] *Resource-Oriented Lightweight Information Exchange (ROLIE), RFC 8322* [Online]. Available: https://tools.ietf.org/html/rfc8322

OpenIOC has been released by "Mandiant" as an Open Source project in November of 2011. It was designed to enable logical comparison of indicators with the help of "AND" and "OR" operators. Using logical operators allows to increase threat description flexibility and improve threat detection rates in comparison with usage of traditional malware signatures [58]. The project has been extended to a framework that allows to manage, search and exchange IOCs at machine speed. Future development of OpenIOC is aiming at providing even more flexibility by improving indicators and supplying IOCs with metadata extensions. OpenIOC framework uses XML language, allowing to create extensible schemas that describe technical characteristics of cyber threats [56], [59]. Example of an OpenIOC is shown on figure 7.

Figure 7. Zeus malware IOC. Source: [59].

Benefits of OpenIOC framework include [56]:

- Free to use under Apache 2 license
- XML language allows to extend IOC descriptions as needed
- OpenIOC software family includes tools to create, modify and search for OpenIOC indicators
- Full support for "Mandiant" products

Disadvantages of OpenIOC framework [56]:

- Limited interoperability with non-Mandian products
- Limited support for working with network-based IOCs
- Complicated integration with IDS
- Insufficient capabilities to describe adversary TTPs

CybOX, MAEC, STIX, and TAXII are community-developed open-source and free to use standards and languages. “Development of these tools has been moderated by the MITRE Corporation, and sponsored by the office of Cybersecurity and Communications and the U.S. Department of Homeland Security”[1]. STIX 1.X and CybOX 2.X versions have been merged into STIX 2.0, which is currently maintained by the OASIS CTI TC[2].

Structured Threat Information Expression (**STIX™**) is a language and serialization format used to exchange TI. STIX information can be visually represented for an analyst or stored as JSON to be quickly machine readable (see figures 8 and 9). All aspects of suspicion, compromise and attribution can be represented clearly with twelve STIX Domain Objects (SDOs) and descriptive STIX Relationship Objects (SROs). STIX's openness allows for integration into existing tools and products or utilized for your specific analyst or network needs [60].

```
{
    "type": "campaign",
    "id": "campaign--8e2e2d2b-17d4-4cbf-938f-98ee46b3cd3f",
    "created": "2016-04-06T20:03:00.000Z",
    "name": "Green Group Attacks Against Finance",
    "description": "Campaign by Green Group against targets in the financial services sector."
}
```

Figure 8. JSON-based example of a STIX 2.0 Campaign object. Source: [60].

[1] CybOX Terms of Use available at https://cybox.mitre.org/about/termsofuse.html

[2] STIX 2.0 documentation is available at Oasis CTI TC Github repository https://oasis-open.github.io/cti-documentation/

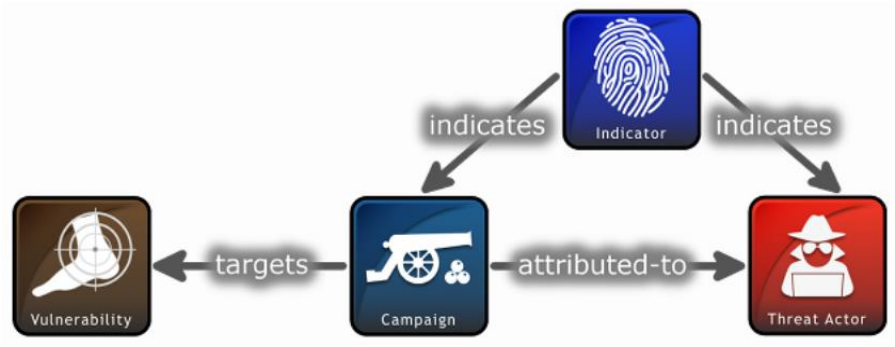


Figure 9. STIX 2.0 relationship example. Source: [60].

Malware Attribute Enumeration and Characterization (MAEC) is "a structured language for encoding and sharing high-fidelity information about malware based upon attributes such as behaviors, artifacts, and relationships between malware samples"[1]. Development of MAEC was inspired by the demand for community-accepted standard that could describe characterization of malware using abstract patterns, instead of physical signatures. Similarly to STIX, MAEC defines several top-level objects and relationships between the objects (including STIX objects, see figure 10), which allows to visualize malware descriptions. JSON schemas allow to feed MAEC data into security tools for automated processing [61].

MAEC developers aim to provide the following major benefits for the community:

"Elimination of ambiguity and inaccuracy in malware descriptions – MAEC should improve human-to-human, human-to-tool, tool-to-tool, and tool-to-human communication about anti-malware related information. This will positively impact all major stakeholders, including producers and consumers of malware analysis and related malware data, as well as the end-users of tools for malware prevention and mitigation [61]".

[1] The MAEC Project documentation is available at Github repository https://maecproject.github.io/documentation/

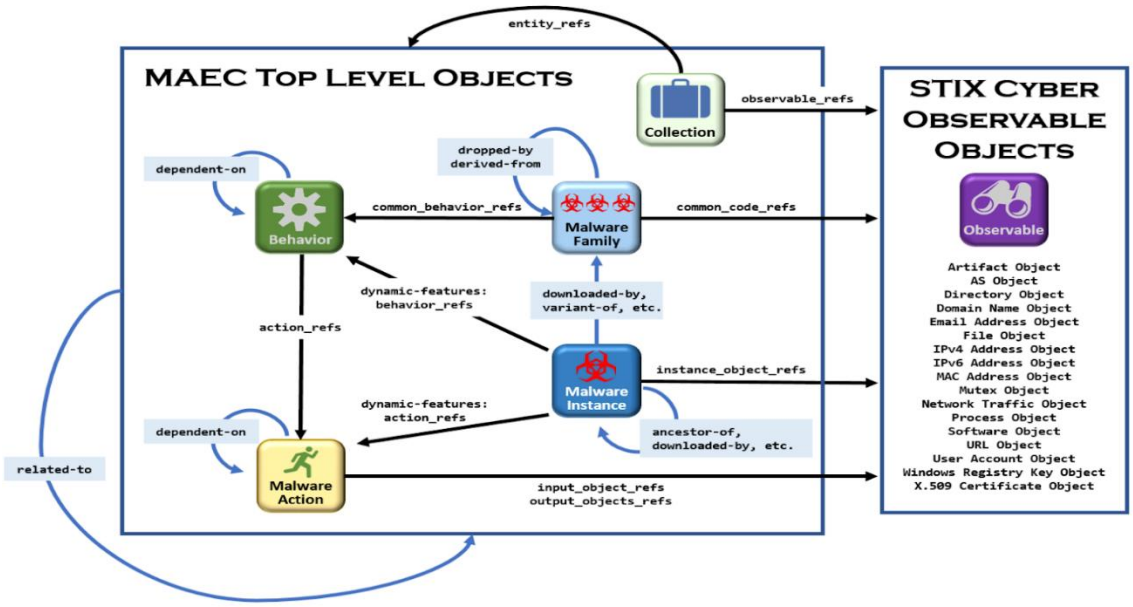


Figure 10. MAEC Top Level Objects and relationships with STIX observables. Source: [61].

- "Reduced duplication of malware analysis efforts – A common method of characterizing malware, along with a corresponding standard for malware analysis reporting, will allow researchers and analysts to determine whether a particular malware instance has already been analyzed [61]".

- "Improved general awareness of malware – A widely adopted standard for characterizing malware will allow for increased public awareness of malware threats and activity [61]".

- "Decreased overall response time to malware threats – MAEC's standard method of describing malware behavior will enable countermeasures for previously observed malware instances to be leveraged, resulting in faster mitigation and response [61]".

Trusted Automated Exchange of Intelligence Information (TAXII™) is an application protocol developed as the preferred exchange mechanism for STIX content. However, TAXII can be used to transport non-STIX data as well. TAXII defines a

RESTful API (a set of services and message exchanges) and a set of requirements for TAXII Clients and Servers[1]. The three principal models for implementing TAXII include:

- "Hub and spoke - one repository of information,
- Source/subscriber - one single source of information,
- Peer-to-peer - multiple groups share information [62]."

TAXII defines four services that can be combined by users to implement different TI sharing models:

1. "Discovery - a way to learn what services an entity supports and how to interact with them,
2. Collection Management - a way to learn about and request subscriptions to data collections,
3. Inbox - a way to receive pushed content (push messaging),
4. Poll - a way to request content (pull messaging) [62]".

3.3.3 Sources of IOCs.

To begin using IOCs, CSIRT has to find a way to initially obtain or develop them. Sources of IOCs can be divided into three main categories: external, community, and internal; each having its benefits and shortcomings [39], [42], [55].

Internal sources. Although a large amount of IOCs can be gathered from various external and community sources, so-called "global IOCs" can be irrelevant to a particular organization. The focus of threat hunting can be switched from collection to developing and enriching organization's specific IOC data. This can be achieved by populating organization's environment with internally discovered IOCs and correlating it with other

[1] TAXII introduction is available at Oasis CTI TC Github repository [Online]. https://oasis-open.github.io/cti-documentation/stix/intro

artifacts found within the environment. Such approach helps to produce context, which would be relevant to a specific threat, facing particular organization [63].

Community sources. Organizations that are members of same industry sector, or that have common interests can establish trusted relationships to exchange IOCs. ISACs have already been mentioned in section 3.1.2. as communities that are formed to facilitate cooperation within specific industry sectors in terms of TI sharing. Specific organizations without established ISACs can participate in TI sharing programs with the help of National Council of ISACs (NCI). NCI coordinates collaboration between individual ISACs and their cooperation with governmental structures to protect facilities, personnel and customers from cyber threats [55], [64].

New examples of malware or attack patterns are often discussed by security community online. In some cases, specific IOC distribution websites and discussion boards may contain information about recently discovered threats and associated indicators. Even though such threat-related information is not always presented in solid and structured manner, description of certain IOCs and artifacts can be obtained from the Internet free of charge [42].

External sources. Threat intelligence can be acquired from sources outside an organization either for a paid basis, or free of charge. Several security vendors provide subscriptions to threat feeds, which are known as *private* external TI sources. Advantage of intelligence obtained from security vendors is its quality (because it is acquired in a more timely manner, compared to other data sources), variety (commercial feeds can provide more detailed information related to specific products, which are released by a particular vendor), and regular updates [55].

4 Creating a tool framework for automating incident triage.

The following part of thesis describes a practical part of author's research on available open-source tools and their capabilities. The aim of this research is to find a solution, which would enhance incident management through partial automation of its sub-processes and specifically incident triage.

Theoretical study of incident management and threat intelligence topics allows to define a list of requirements regarding a possible solution for partial automation of triage of cyber incidents. This chapter describes author's approach in creating a framework of open-source tools, which would meet the defined requirements. Chapter also includes a description of case study conducted by author to test the proposed framework and determine its efficiency.

4.1 Framework requirements and architecture.

Before starting the analysis of possible tools for the proposed framework, author defined three general requirements, which would have to be satisfied by every potential component. Firstly, the tools would have to be scalable in terms of being able to handle different amount of processed data. Such requirement is considered by author as an assurance of framework's suitability for organizations with small networks, and companies with more complex information system infrastructure. Secondly, the tools would have to provide full control of processed system-, and incident-related data to reduce risks related to disclosure of sensitive or classified information to third parties. And lastly, the tools would need to have developed APIs to support their integration into the framework, as well as automation of related workflows. Additional requirements to individual framework elements have been defined in accordance with their functionality features.

As author declared in section 1.2., the scope of this research is limited to triaging cyber incidents, which can be detected through hunting for IOCs within a given environment. Thus, a potential framework should have included a tool, which would allow to create, collect and store IOCs in a single repository. Malware Information Sharing Platform ("MISP") has been used in the proposed framework for such purpose, because:

- “MISP” has a built-in sharing functionality, which makes it easy to participate in exchange of threat-related information with other parties and consume IOCs;
- On the other hand, “MISP” allows to store IOC-related information locally and ensure confidentiality of this data, if necessary;
- “MISP” supports classification of processed IOCs in accordance with “eCSIRT.net” taxonomy, which is important in terms of NIS directive (and particularly important for exchanging threat- and incident-related information in an internationally accepted and structured manner);
- Questioning of CSIRTs with national responsibilities showed that “MISP” is widely used for exchanging IOCs in practice;
- “MISP” provides a certain level of automation when used in combination with other framework tools.

To hunt organization’s environment for particular IOCs and detect corresponding incidents (including those, which could have occurred before a particular IOC was obtained via MISP), a full packet capturing (FPC) tool should have been included into the framework. Considering framework’s general requirements, “Moloch” has been chosen as a FPC system, since it:

- Can scale to handle tens of gigabits/sec of network traffic;
- Allows to store and index network traffic in standard PCAP format;
- Provides fast access to the stored data via it’s API;
- Provides complete control over the processed data.

Finally, the tool framework should have included an incident tracking repository, which would allow to register detected incidents and save any relevant information for future analysis. For this purposes author has chosen “TheHive” platform because:

- It is designed to be easily synchronized with one or multiple “MISP” instances and automatically receive events, published in the “MISP”;

- it has a built-in functionality, which allows to receive “MISP” tags describing classification of cyber threats;

- It has an ability to automatically correlate data by identifying IOCs that have been already associated with previous incidents;

- Although this feature is outside of the scope of this research, “TheHive” allows to enhance analysis of incident-related IOCs with the help of “Cortex” analyzers. Such advantage is implemented as a single point of querying, allowing for automation via “Cortex” REST API.

The described set of tools has been tested by the author during practical part of this research. The details of a conducted case study are described in the following chapter.

4.2 Case study.

4.2.1 Description of test environment.

To conduct a practical research author prepared a virtual network, which included 3 subnets separated into different VLANs (see appendix 1). Such topology provided just enough flexibility in terms of estimating criticality of hosts, located in different segments of the prepared environment.

“LAN” subnet consisted of 4 VMs, configured with 2 cores and 4 GB of memory, each running “Windows 7” operating system. The workstations were accessible over RDP connections and simulated an intranet of an organization.

“DMZ” subnet included 4 VMs, configured with 2 cores and 4 GB of memory, operated by Debian “jessie” operating systems and accessible over SSH. “DMZ” represented a network segment, which is exposed to the Internet.

“MGMT” subnet consisted of 4 VMs operated by Debian “jessie” servers and accessible over SSH. These servers have been set up with different parameters and used by author for deploying and testing interoperability and efficiency of selected tools.

Virtual switch has been configured to mirror network traffic (both ingress and egress) via a SPAN session, where “LAN” and “DMZ” were configured as source VLANs, while one

of the hosts in “MGMT” subnet has been set up with 2 network interfaces to be capable of receiving the mirrored traffic.

“MISP” instance (v 2.4.8.8) has been installed on one of the Debian servers (10.101.3.106), configured with 1 core and 512 MB of memory.

To store and access network traffic, the last stable version of “Moloch” (v0.20.2 at the moment of conducting this research) has been installed on another server (10.101.3.105) together with the latest version of “Elasticsearch” (v 5.6.8) supported by this particular version of “Moloch”. This VM has been configured with 4 cores and 8 GB of memory.

The third Debian server (10.101.3.4) has been used to install the following “TheHive” components:

- TheHive 3.0.6
- Elastic4Play 1.4.5
- Play 2.6.7
- Elastic4s 5.6.0
- ElasticSearch 5.6.2

“TheHive” has been synchronized with the “MISP” instance and configured to check for new events every 5 minutes. The host itself has been configured with 2 cores and 4 GB of memory.

4.2.2 Testing the framework.

Proposed framework testing occurred in several steps. First, the author used hosts belonging to “LAN” and “DMZ” subnets to simulate malicious activity within the test environment. Simulation has been performed by downloading sample files onto randomly selected hosts belonging to “LAN” and “DMZ” subnets. The “malicious content” has been downloaded from the following locations:

- www.fakefilegenerator.com
- http://ipv4.download.thinkbroadband.com/5MB.zip

- http://212.183.159.230/10MB.zip

The next step performed by author was to create and publish “MISP” events, which would describe different attributes of the files (IOCs), downloaded during previous step (see figures 14 – 16). In addition to attribute descriptions each event included a “MISP” tag, which described the class of a particular threat in accordance with “eCSIRT.net” taxonomy as follows:

- Important_Document.pdf – “malware”;

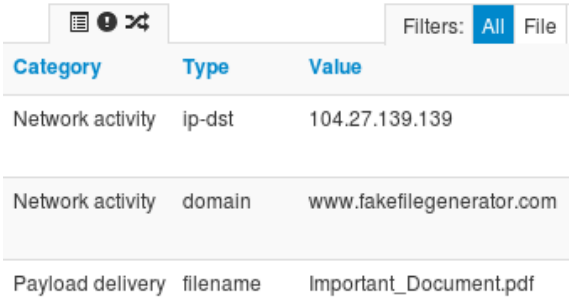


Figure 11. A list of MISP attributes related to „Important_Document.pdf“.

- 5MB.zip – "ransomware";

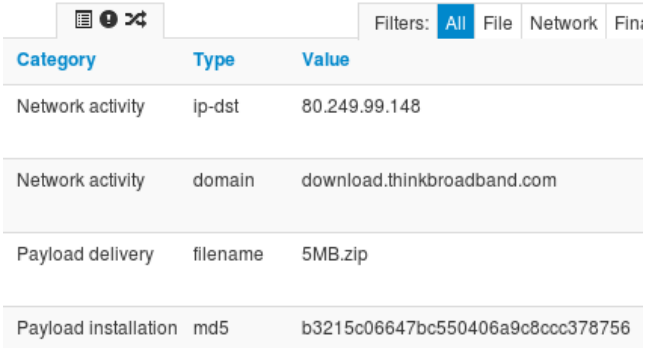


Figure 12. A list of MISP attributes related to „5MB.zip“.

- 10MB.zip – “trojan”.

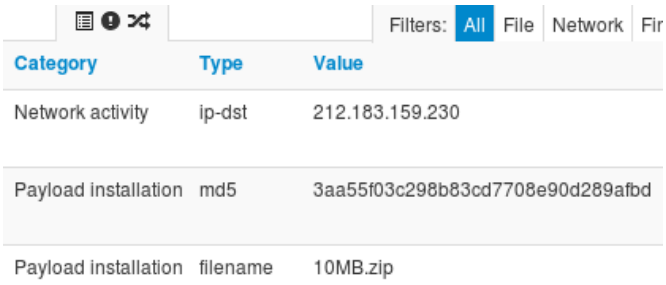


Figure 13. A list of MISP attributes related to „10MB.zip“.

After “MISP” events, which described downloaded “malicious files” were published, “TheHive” automatically imported them and displayed under “Alerts” panel (see figure 17).

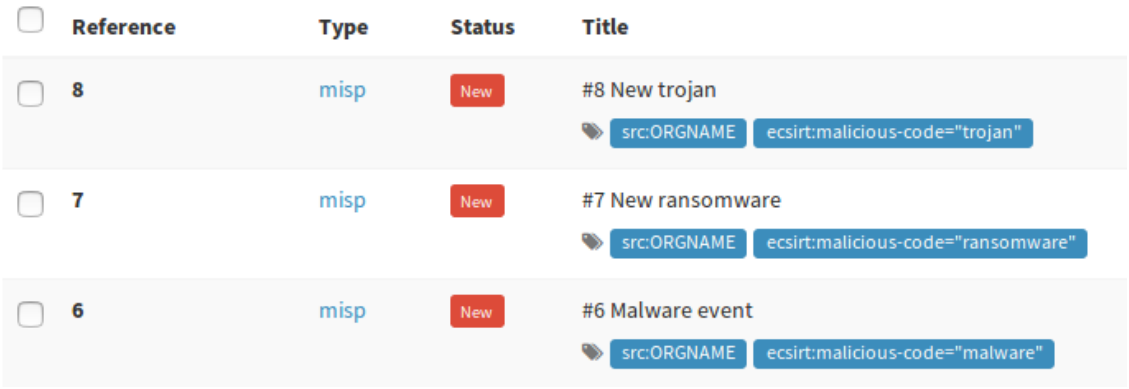


Figure 14. „TheHive“ alerts, created from published „MISP“ events.

Each alert imported by “TheHive” represented a description of a potential threat to the given test environment. As described in section 2.2., triage begins with incident-related information categorization and verification of security events.

At the moment, when “TheHive” imports “MISP” events, their description already includes the classification of a given threat, allowing to identify adversary’s attack methods. To verify a security event author had to detect attack targets in addition to the attack methods [15]. For this purpose author used IOCs obtained from “TheHive” alerts to search network traffic for matches, which would allow to determine potentially affected hosts. IOC hunt has been performed with the help of a script, created by the author. At this stage the script performed the following actions:

- queried “TheHive” API to check, if any alert is marked with the status “New”;

- if “New” alerts could be found, the script used alert’s unique ID to parse the data stored in the alert and extract:
    - the “MISP” tag, which has been used to classify a particular threat in accordance with eCSIRT.net taxonomy;
    - available IOC types (e.g. IP addresses, domain names) and their corresponding values;
- used types and values of extracted IOCs to compile an API request towards “Moloch”;
- queried the “Moloch” via it’s API to search network traffic for extracted IOCs and determine IP addresses of potentially affected hosts within the test environment.

After this part of the script has been executed, information gathered from the test environment allowed to continue with the next triage sub process – incident prioritization. Author calculated incident priority by multiplying corresponding values of *threat severity* [27] and potentially affected *host importance*. *Host importance* could be estimated using three different methods:

1. calculated using the Euclidian norm of host’s confidentiality, integrity and availability metrics as proposed in [30, 36];
2. Using the highest determined value among host’s confidentiality, integrity and availability metrics as proposed in [37];
3. Calculated in accordance with the network security zone to which the host belongs as proposed in [30 - 32].

The first and the second methods would provide more accurate results, because this would allow to confront host metrics with a specific vulnerability or threat features. However, automated estimation of individual host confidentiality, integrity, and availability metrics in large environments, or networks with frequent changes in workstation numbers can become a challenging task.

This is why author decided to evaluate host importance in accordance with [30 - 32]. Such method of *host importance* estimation allows to automatically assign corresponding

values to individual hosts whenever they obtain IP addresses from predefined ranges of specific network segments. Accuracy of the selected method can be considered insufficient, since it doesn't allow to take into account host confidentiality, integrity and availability requirements. However, since selected method relies on IP addresses, it allows to define individual importance values for hosts that might require specific attention. This can be performed independently of *host importance* values assigned to other hosts belonging to the same security zone. Thus, accuracy of the selected method can be improved through a more detailed documentation of network topology.

To make *threat severity* and *host importance* values automatically accessible for the script, author has prepared 2 files on the same host, where the script was executed.

Content of the first file represented results of a simplified measurement of threats relevant to the environment. A list of such threats should have been provided to a CSIRT as a result of threat assessment conducted by organization's management. For the purpose of research, author measured every threat, which is included into "eCSIRT.net" taxonomy (see figure 18).

Author used numbers from 1 to 10 to assign each threat with a respective severity value, where 10 represented the highest severity for a given threat. It would be important to mention that such scale, as well as assigned threat severity values should not be taken for granted. Such model has rather been used as an example and can be adjusted to meet specific requirements of any organization.

The second file contained information about organization's network topology / security zones, corresponding IP address ranges and zone importance values (see figure 19). Similarly to the contents of the first file, information stored in the second file has been used as an example and should be changed to reflect actual situation within a specific environment.

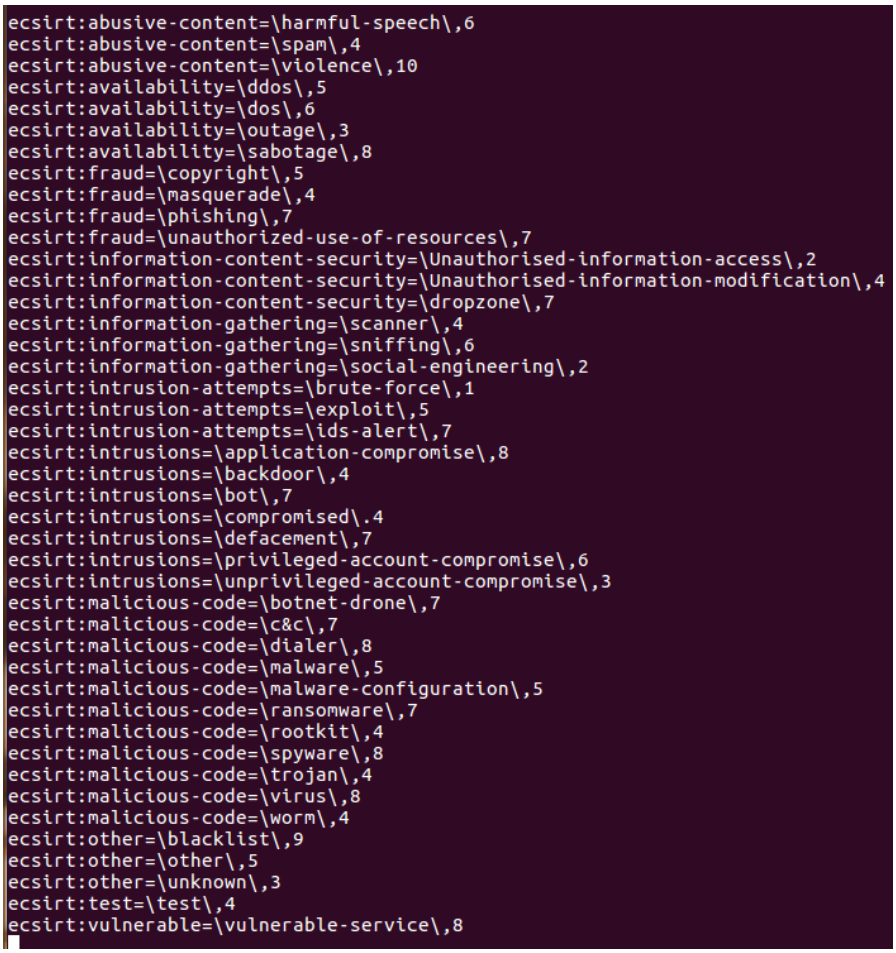


Figure 15. An example of threat measurement results, stored in a local file.

A list of security zones describing network topology can additionally include IP addresses of unique hosts, whenever some critical hosts require specific attention in terms of prioritizing incidents. Such changes in contents of this file would require the script to be adjusted accordingly, so it would be able to detect precise address matches, instead of determining, which security zone affected host belongs to.

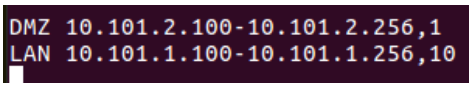


Figure 16. An example of security zones description, stored in a local file.

The script itself has been designed to associate each potentially affected host with an individual incident. To estimate each individual incident's priority the script:

- Referred to the file containing threat measurement results and searched its lines for a match with the value of the tag (describing class of a particular threat)

extracted from “TheHive” alert. Upon match detection assigned a particular threat with corresponding *threat severity* value.

- Referred to the file containing description of security zones and used IP addresses of potentially affected hosts to determine, which security zone they belong to. Assigned individual hosts with a *host priority* value in accordance with their location within a test environment.

- Multiplied *threat severity* and *host priority* values to calculate *incident priority*.

In addition to triaging individual incidents, the script followed ITIL principle [26] to estimate overall impact, which is potentially caused to the test environment by a particular threat. This has been performed through consolidation of the individual incident priority values.

After individual incidents have been verified, classified, prioritized and overall impact of a particular threat estimated, the script passed the obtained / calculated data to “TheHive”. During this step author determined a shortcoming of “TheHive” platform. Author’s approach for estimating incident priority uses a flexible scale to improve accuracy of the results. However, “TheHive” design is limited to only accept 3 possible incident severity values: “Low”, “Medium” and “High”. To bypass such limitation, author adjusted the script to display calculated incident priority in the name of corresponding case, created by “TheHive”. In order to register determined incidents and pass relevant data to “TheHive”, the script followed an algorithm, described below:

- Created a new case for each host, which has been detected as potentially affected;
- Added a descriptive tag identifying the class of a threat, which has caused the incident;
- Reflected a “MISP” event name used to detect incidents together with the calculated priority value of individual incident in a title of created case;
- Used a “description” field to provide additional information related to the incident, including IP address of affected host and a name of security zone, in which the incident has been detected (see figure 20);

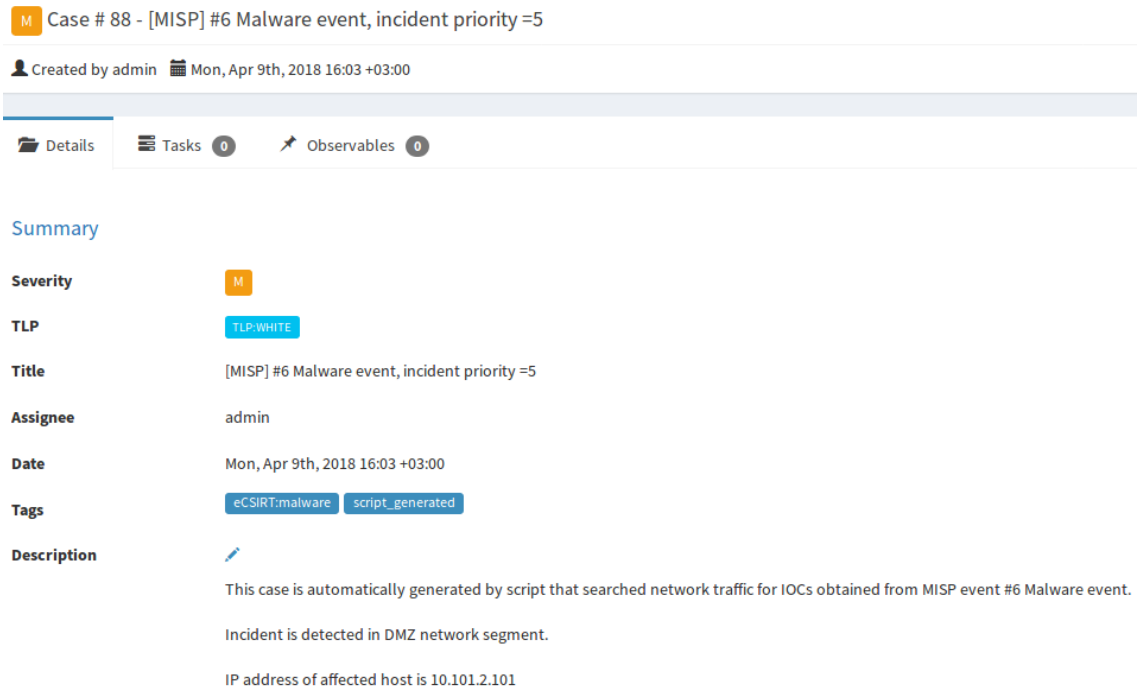


Figure 17. A case created in „TheHive“, describing individual incident.

After the script has created a case for each individual incidents, it created an additional “parent” case to consolidate information related to a particular threat impact on the test environment (see figure 21). “Parent” cases were created by the script in accordance with the following algorithm:

- A new case was created for each alarm, imported from the “MISP”;
- The script added a descriptive tag identifying the class of a threat affecting the test environment;
- The script identified the case as “parent” in its title;
- The script reflected a “MISP” event name used to detect incidents together with the value of estimated overall impact of a threat on the test environment in a title of a created case;
- The script used a “description” field to provide additional information related to the incident, including a total amount of hosts, which were potentially affected by a particular threat, and their IP addresses.

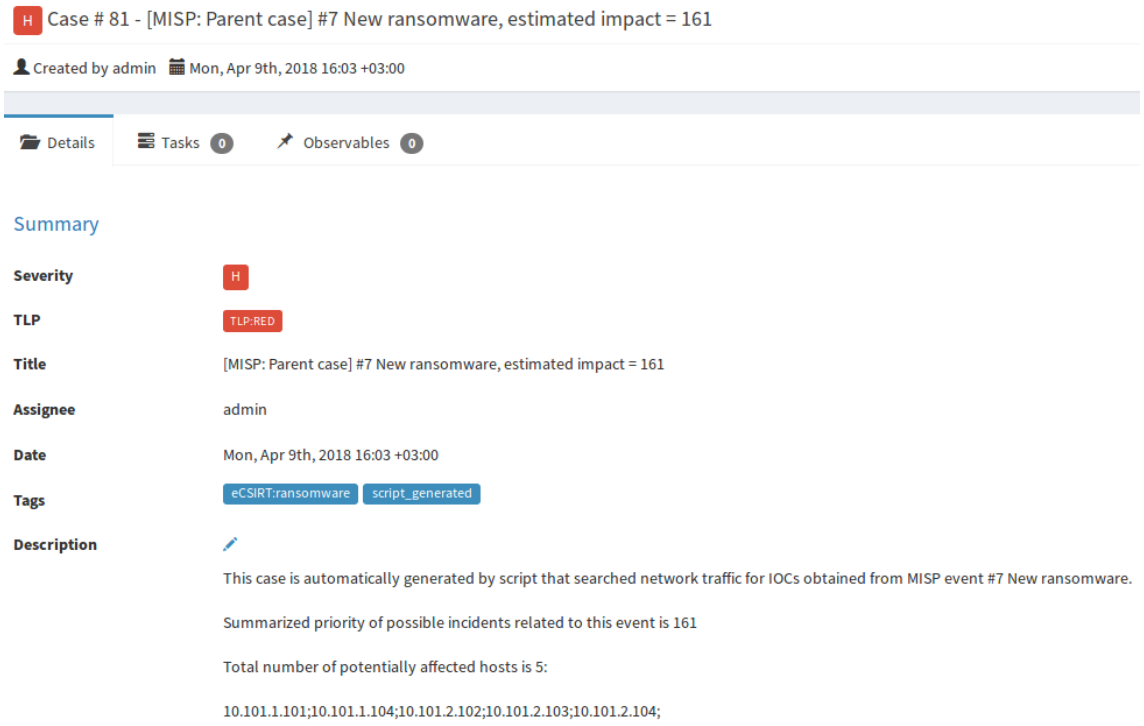


Figure 18. A „parent“ case created in „TheHive“, describing overall impact of a particular threat.

During evaluation of proposed framework, the script operated with 3 different “TheHive” alarms, imported from the “MISP”. Files, which simulated malicious content have been downloaded onto randomly selected host located in different security zones. The amount of hosts, which were used to download “malicious content” varied depending on the type of an examined threat. Such approach has been selected by the author to illustrate how estimated value of overall threat impact would change depending on the total amount of potentially affected hosts.

The lowest priority (estimated impact = 55) was assigned to a “parent” incident, which aggregated information about hosts affected by “malware”. One of the potentially infected hosts was located in “DMZ” (corresponding incident priority = 5), and another in “LAN” (corresponding incident priority = 10) security zones.

A “parent” incident, describing impact of a “new ransomware” was assigned the highest priority (estimated impact = 161). The script verified three respective incidents in “DMZ” security zone (each assigned a priority value of 7), and two incidents in “LAN” (both with priority value of 70).

Four individual incidents related to a “new trojan” were detected during case study. Three of them were verified in “LAN” subnet (each assigned a priority value of 40), and one affected a host, located in “DMZ” (assigned a priority value of 4). The script estimated overall impact of a “new trojan” as 124 and created a respective “parent” case.

The list of “TheHive” cases created as a result of script execution is shown on figure 22.

Figure 19. The list of „TheHive“ cases created by a script during case study.

5 Conclusion.

This chapter summarizes the thesis in section 5.1. by evaluating, whether selected research methods and results of the conducted case study allowed to answer the main question of the research. Section 5.2. includes discussion about author's contribution, limitations of the practical research and capabilities of the proposed solution. In section 5.3. author discusses several possibilities to improve framework's efficiency during future research. Finally, section 5.4. concludes the thesis.

5.1 Thesis summary.

Evolution of information systems technologies and nature of cybercrime creates various challenges for cyber security professionals. Continuous increase in complexity of information systems results in enormous amount of security-related data, which needs to be processed by CSIRTs. Effectiveness of CSIRTs can be increased through automating workflows, but it typically assumes that CSIRTs have enough resources to invest into specific security products, or expert intelligence. However, CSIRTs are commonly limited in their resources and typically can't afford themselves to make such investments.

Such situation places attackers of information systems in a beneficial position. While security professionals attempt to guard every possible attack vector, and get overwhelmed by the data originated from various sensors, attackers only have to exploit a single vulnerability to begin infiltrating into the protected systems.

Participating in exchange of threat intelligence allows CSIRTs to keep the pace with adversaries and improve knowledge about attackers' tactics, techniques and procedures. Such knowledge helps to decrease chances of successful attacks, or detect them earlier in the kill chain and prevent adversary intrusion.

Nevertheless, every attack executed against protected systems cannot be prevented. Incidents will still occur and CSIRTs should be able to manage them adequately. Incident triage is a process performed by CSIRTs to distinguish (verify) security incidents from other triggered alarms, classify verified incidents and assign them priority values. Many factors affect a selection of possible triage implementation model. However, despite of

the chosen triage implementation approach, it should be perceived as an indispensable process for effective incident handling.

The main purpose of this research was to develop an open-source tool framework, which would allow to automate triage of the incidents, detected through searching for technical indicators of compromise within a given environment. Such solution would allow CSIRTs with limited budget to increase their efficiency through partial automation of their workflows and improved incident detection rates.

Author began the research with theoretical study of incident management and threat intelligence sharing topics. The purpose of this work was to:

- determine the role of cyber incident triage and its relations with other incident management sub processes;
- define incident triage sub-processes and methodologies of their implementation;
- analyze existing approaches in estimating severity of cyber incidents;
- describe challenges related to threat intelligence consumption and possible methods of their overcoming;
- analyze benefits and methodologies of using threat intelligence and specifically indicators of compromise;
- compare existing types of indicators of compromise and possible methods to improve their quality.

Chapter 4 describes a case study, conducted by author practically. Tested open-source tool framework included three software products: “MISP”, “Moloch” and “The Hive”. The tools were selected, based on the requirements determined during theoretical study and their interoperability capabilities. Efficiency of a tool framework was tested in a virtual environment by simulating malware-based cyber incidents and subsequent hunt of respective indicators of compromise.

Conducted case study confirmed that tested framework is capable of verifying cyber incidents with the help of indicators of compromise included into published in “MISP” events. To classify verified incidents author used the “MISP” tagging feature and its capability to describe published events in accordance with a wide range of available taxonomies. Priority of incidents was estimated considering severity of a threat affecting

the environment, and host importance, which was determined by criticality of a security zone that the host belongs to. Automation of interaction between framework elements was achieved with the help of a script written by author. In result of the script execution, it registered verified incidents as "The Hive" cases, enriching them with incident-relevant information obtained during triage process.

5.2 Discussion.

5.2.1 Author's contribution.

The proposed tool framework exclusively includes open-source and free-to-use security software products. Practical research, conducted by author confirmed that the framework allows to enhance incident management through partial automation of its sub processes. This thesis allows CSIRTs to evaluate suitability of framework for their IT infrastructure, and apply the designed framework without a demand for additional financial investments.

To automate interaction between components of the proposed framework, author developed a Bash script, which takes advantage of selected tools' APIs. The script allows to use threat-related information obtained from the "MISP" for automated detection of cyber incidents within a given environment, and their subsequent triaging in accordance with the proposed method.

Theoretical study of threat assessment topic and existing cyber incident scoring approaches allowed author to design a scalable method to estimate priority of incidents, based on threat severity and potentially affected host importance values. Simplicity of the proposed method, compared to possible alternative approaches analyzed by author, allows to utilize it in different environments, regardless of their complexity.

5.2.2 Limitations of the thesis.

Even though conducted case study confirmed that the proposed solution allows to partially automate incident detection and triage processes, author is aware of several limitations in regards to study design and capabilities of the proposed framework.

Study design limitations. Cyber incidents can be detected with the help of security-related information, obtained from a wide range of data sources. For the purpose of this

research, author limited his scope to automating triage of incidents, which can be detected by performing IOC checks within a given environment.

Taxonomy proposed by “eCSIRT.net” and used in this research includes 43 different classes of cyber threats. Author focused on designing algorithm for performing triage of three specific classes of incidents, which can be referred to as “malware-based”.

Limitations of proposed solution. Malicious files can be described with the IOCs, which combine host-, network-, and malware-based artifacts. Developed script operates using network-based artifacts, obtained from published “MISP” events. This allows the script to detect potentially affected hosts, but limits its sufficiency in regards to security event correlation capabilities.

Proposed framework was designed to analyze network traffic, which is stored before the script begins processing alerts generated from “MISP” events. This limits the value of events published in the “MISP” in context of developed algorithm, since conducted research didn’t include implementation of their potential for detecting incidents by inspecting future traffic flowing through the environment.

5.3 Future work.

Taking into consideration the limitations of study design and proposed solution capabilities, efficiency of the designed framework can be extended in several directions.

Improving incident prioritization accuracy.

Malware-based incidents detected in practical part of this research have been prioritized without consideration of predisposing conditions. Such incidents were triaged as if malware, which has been downloaded onto the virtual hosts was executed afterwards. However, if security tools would have prevented malware from execution, or deleted the malicious content immediately after it was downloaded, priority of respective incidents should have been adjusted accordingly. Proposed framework can be expanded with additional software components, which would allow to perform correlation of security events, related to host-, and malware based artifacts. This would increase effectiveness of

using IOCs and enhance incident triage process by providing incident managers with more accurate information in regards to estimated incident priority.

Detecting and triaging future incidents.

Execution of the script allows to search stored network traffic for the presence of indicators of compromise, obtained from the published "MISP" events. This allows to detect and triage unknown incidents, which could have occurred before the script was executed. However, potential of IOCs allows to use them for detection of malicious activity after implementation of IOCs in a given environment. Thus, efficiency of the proposed framework can be amplified by developing a complete solution in terms of incident detection. This would allow to perform triage of cyber incidents, independently of "The Hive" alert processing timeframe.

Expansion of "TheHive" alert sources.

For the purpose of conducting research, scope of this thesis was limited to triaging incidents, which can be detected with the help of hunting stored network traffic for the presence of indicators of compromise, obtained from published "MISP" events. However, developers of "TheHive" platform claim that its Python API client is capable to aggregate security events from other sources and use them to create new alerts. A list of sources of precursors and indicators published in [11] can be used to expand a list of "TheHive" alerts generators and broaden capabilities of the proposed framework.

References

[1] B. M. Leiner, V. G. Cerf, D. D. Clark, R. E. Kahn, L. Kleinrock, D. C. Lynch, J. Postel, L. G. Roberts, S. Wolff. (1997) "Brief History of the Internet", Internet Society [Online]. Available: https://www.internetsociety.org/internet/history-internet/brief-history-internet/

[2] H. Orman. (2003, Sept.-Oct.) "The Morris Worm:A Fifteen-Year Perspective." IEEE Security & Privacy [Online]. vol. 99, issue 5. Available: http://ieeexplore.ieee.org/document/1236233/

[3] Bill Horne. (2014, Sept.-Oct.) "On Computer Security Incident Response Teams." IEEE Security & Privacy [Online]. vol. 12, issue 5, pp 13 - 15. Available: http://ieeexplore.ieee.org/document/6924687/

[4] X. Li, (2017) "A Review of Motivations of Illegal Cyber Activities", Criminology & Social Integration Journal [Online]. Vol. 25 No. 1. Available: https://hrcak.srce.hr/file/266976

[5] R. M. Lee, "The Sliding Scale of Cyber Security," a SANS Analyst Whitepaper [Online], Aug 2015. Available: https://www.sans.org/reading-room/whitepapers/analyst/sliding-scale-cyber-security-36240

[6] K. Chang. (2014, Jun. 2) Automating Cybersecurity [Online]. Available: https://www.nytimes.com/2014/06/03/science/automating-cybersecurity.html

[7] T. Cole, "Diagnosis SOC-Atrophy: What To Do When Your SOC Is Sick," in RSA Conference, San Francisco, 2017, Feb 13 – 17 [Online]. Available: https://www.rsaconference.com/writable/presentations/file_upload/air-w11-diagnosis_soc-atrophy-_what-to-do-when-your-soc-is-sick.pdf

[8] S. Bhatt, P. K. Manadhata, L. Zomlot. (2014, Oct. 15) "The Operational Role of Security Information and Event Management Systems." IEEE Security &

Privacy [Online]. vol. 12, issue 5, pp 35 - 41. Available: http://ieeexplore.ieee.org/document/6924640/

[9] A. Torres, "Automation in the Incident Response Process: Creating an Effective Long-Term Plan," a SANS Whitepaper [Online], Feb. 2015. Available: https://www.sans.org/reading-room/whitepapers/analyst/automation-incident-response-process-creating-effective-long-term-plan-35802

[10] "Cyber Security Incident Management Guide", Cyber Security Coalition White paper, 2016 [Online]. Available: https://www.cybersecuritycoalition.be/content/uploads/cybersecurity-incident-management-guide-EN.pdf

[11] P. R. Cichonski, T. Millar, T. Grance, K. Scarfone, "Computer Security Incident Handling Guide," Recommendations of the National Institute of Standards and Technology [Online], Aug. 06, 2012. Available: http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf

[12] C. Alberts, A. Dorofee, G. Killcrece, R. Ruefle, M. Zajicek, "Defining Incident Management Processes for CSIRTs: A Work in Progress", CMU/SEI, Pittsburgh, TR-015, 2004. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.207.8190&rep=rep1&type=pdf

[13] "Incident Management and Response", ISACA White paper, Mar. 2012 [Online]. Available: http://www.isaca.org/Knowledge-Center/Research/Documents/Incident-Management-and-Response_whp_Eng_0312.pdf

[14] "Incident Handler's Handbook", P. Kral, The SANS Institute, 2012 [Online]. Available: https://www.sans.org/reading-room/whitepapers/incident/incident-handlers-handbook-33901

[15] “Good practice guide for incident management,” M. Maj MSc, R. Reijers, D. Stikvoort MSc, ENISA White paper [Online], Dec. 20, 2010. Available: https://www.enisa.europa.eu/publications/good-practice-guide-for-incident-management

[16] M. J. Wes-Brown, D. Stikvoort, K.-P. Kossakowski, G. Killcrece, R. Ruefle, M. Zajicek, “Handbook for Computer Security Incident Response Teams (CSIRTs)” CMU/SEI, Pittsburgh, HB-002, 2003. [Online]. Available: https://resources.sei.cmu.edu/asset_files/Handbook/2003_002_001_14102.pdf

[17] “Cyber Security Incident Response Guide” Version 1, CREST White paper, 2013 [Online]. Available: https://www.crest-approved.org/wp-content/uploads/2014/11/CSIR-Procurement-Guide.pdf

[18] “Global information Assurance Certification Paper”, N. N. Corothers, The SANS Institute, 2000 - 2002 [Online]. Available: https://www.giac.org/paper/gsec/2022/vulnerability-assessments-methodologies-perform-self-assessment/103498

[19] “Incident Tracking In The Enterprise”, J. Hall, The SANS Institute White paper, 2015 [Online]. Available: https://www.sans.org/reading-room/whitepapers/incident/incident-tracking-enterprise-36092

[20] “Incident Categories (Public)”, CERT gouvernemental Luxembourg, White paper, Jan 2016, [Online]. Available: https://www.govcert.lu/docs/PRO303_Incident_Categories_(Public)_3.0.pdf

[21] “Incident Classification / Incident Taxonomy”, eCSIRT.net, J. Arvidsson, D. Stikvoort, [Online]. Available: https://www.trusted-introducer.org/Incident-Classification-Taxonomy.pdf

[22] S.D. Applegate, A. Stavrou, K. Podins, J. Stinissen, M. Maybaum, "Towards a Cyber Conflict Taxonomy", presented at the 5th International Conference on Cyber Conflict, 2013 [Online]. Available: https://ccdcoe.org/sites/default/files/multimedia/pdf/d3r1s2_applegate.pdf

[23] "A good practice guide of using taxonomies in incident prevention and detection", ENISA White paper, Dec 2016 [Online]. Available: https://www.enisa.europa.eu/publications/using-taxonomies-in-incident-prevention-detection

[24] "Guide for Conducting Risk Assessments", NIST Special Publication 800-30, Revision 1, [Online], Sep 2012. Available: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf

[25] "Virginia Tech Guide for Cyber Security Incident Response", The University of Virgina White paper, Version 5.0 [Online], 2016. Available: https://security.vt.edu/content/dam/security_vt_edu/downloads/incident_response.pdf

[26] BMC Software, Inc. ITIL Processes & Best Practices: Incident Management [Online]. Available: http://www.bmc.com/guides/itil-incident-management.html

[27] "An Overview of Threat and Risk Assessment", J. Bayne, The SANS Institute White paper, 202, [Online]. Available: https://www.sans.org/reading-room/whitepapers/auditing/overview-threat-risk-assessment-76

[28] "Vulnerability Assessment", S. Cima, The SANS Institute White paper, Jul 2001, Version 1.2e [Online]. Available: https://www.sans.org/reading-room/whitepapers/basics/vulnerability-assessment-421

[29] "Good Pracice Guide on Vulnerability Disclosure. Fram challenges to recommendations", ENISA White paper, Nov 2015, [Online]. Available: https://www.enisa.europa.eu/publications/vulnerability-disclosure

[30] "Common Vulnerability Scoring System v3.0: Specification Document", FIRST.Org, Inc, [Online]. Available: https://www.first.org/cvss/cvss-v30-specification-v1.8.pdf

[31] A. Kim, M. H. Kang, J. Z. Luo, A. Velazquez, "A Framework for Event Prioritization in Cyber Network Defense", Naval Research Laboratiry, Washington, DC, NRL/MR/5540--14-9541, 2014, [Online]. Available: www.dtic.mil/docs/citations/ADA608707

[32] Homeland Security. (2018, Feb 20). NCCIC Cyber Incident Scoring Systems [Online]. Available: https://www.us-cert.gov/sites/default/files/publications/NCCIC_Cyber_Incident_Scoring_System.pdf

[33] "Infrastructure Security Architecture for Effective Security Monitoring", L. Obregon, The SANS Institute White paper, Dec 2015, [Online]. Available: https://www.sans.org/reading-room/whitepapers/bestprac/infrastructure-security-architecture-effective-security-monitoring-36512

[34] *Information Classification Policy*, ISO/IEC 27001:2005 A.7.2.1, [Online]. Available: http://www.iso27001security.com/ISO27k_Model_policy_on_information_classification.pdf

[35] "Understanding Personally Identifiable Information", I.Goddjin, RiskBased Security [Online]. Available: https://www.riskbasedsecurity.com/reports/RBS-UnderstandingPersonallyIdentifiableInformation-Sept2013.pdf

[36] "HIPAA Basics for Providers: Privacy, Security, and Breach Notification Rules", Department of Health and Human Services, Centers for Medicare & Medicaid Services, Medicare Learning Network, Aug 2016, [Online]. Available: https://www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNProducts/Downloads/HIPAAPrivacyandSecurity.pdf

[37] Information System Authority Estonia. Three-level it baseline security system ISKE, implementation guide, Version 8.00, Jan 2017. [Online]. Available: https://iske.ria.ee/8_03/?action=AttachFile&do=get&target=ISKE%20rakendusjuhend%20ver.%208.00.pdf

[38] "Threat Intelligence: Collecting, Analysing, evaluating", D. Chimson, M. Ruks, MWR Info Security Ltd White paper, 2015, [Online]. Available: https://www.ncsc.gov.uk/content/files/protected_files/guidance_files/MWR_Threat_Intelligence_whitepaper-2015.pdf

[39] "Threat Intelligence: What It Is, and How to Use It Effectively", M. Bromiley, The SANS Institute White paper, Sep 2016, [Online]. Available: https://www.sans.org/reading-room/whitepapers/analyst/threat-intelligence-is-effectively-37282

[40] "Detecting and Preventing Attacks Earlier in the Kill Chain", C. Velazquez, The SANS Institute White paper, Aug 2015, [Online]. Available: https://www.sans.org/reading-room/whitepapers/infosec/detecting-preventing-attacks-earlier-kill-chain-36230

[41] C. Johnson, L. Badger, D. Waltermire, J. Snyder, C. Skorupka, "Guide to Cyber Threat Information Sharing" (Second Draft), National Institute of Standards and Technology Special Publication (SP) 800-150, [Online], Apr. 21, 2016. Available: https://csrc.nist.gov/csrc/media/publications/sp/800-150/archive/2016-04-21/documents/sp800_150_second_draft.pdf

[42] J. Andress. “Working with Indicators of Compromise.” *ISSA Journal.* pp 14-20, May. 2015. [Online]. Available: https://c.ymcdn.com/sites/www.issa.org/resource/resmgr/journalpdfs/feature0515.pdf

[43] C. Beek, D. Frosst, P. Greve, Y. Gund, F. Moreno, E. Peterson, C. Schmugar, R. Simon, D. Sommer, B. Sun, R. Tiwari, V. Weafer, “McAfee Labs Threats Report”, Apr. 2017. [Online]. Available: https://www.mcafee.com/us/resources/reports/rp-quarterly-threats-mar-2017.pdf

[44] “Automated Defence - Using Threat Intelligence to Augment”, P. Poputa-Clean, The SANS Institute White paper, Jan 2015, [Online]. Available: https://www.sans.org/reading-room/whitepapers/threats/automated-defense-threat-intelligence-augment-35692

[45] J. Friedman, M. Bouchard, “Definitive Guide to Cyber Threat Intelligence”, CyberEdge Group, LLC. [Online], 2015. Available: https://cryptome.org/2015/09/cti-guide.pdf

[46] “Comment Crew: Indicators of Compromise”, Security Response, Symantec, Feb 2013, [Online]. Available: https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/comment_crew_indicators_of_compromise.pdf

[47] “Diplomats in Eastern Europe bitten by a Turla mosquito”, ESET, spol. s r.o. Jan 2018, [Online]. Available: https://www.welivesecurity.com/wp-content/uploads/2018/01/ESET_Turla_Mosquito.pdf

[48] “Darkhotel Indicators of Compromise”, Version 1.1, Kaspersky Global Research and Analysis Team, Nov. 2014, [Online]. Available: https://securelist.com/files/2014/11/darkhotelappendixindicators_kl.pdf

[49] M. del C. P. Tixteco, L. P. Tixteco, G. S. Perez, L. K. Toscano Medina, "Intrusion Detection Using Indicators of Compromise Based on Best Practices and Windows Event logs", in *ICIMP 2016: The Eleventh International Conference on Internet Monitoring and Protection*, Valencia, Spain, 2016, pp 29-37, [Online]. Available: www.thinkmind.org/download.php?articleid=icimp_2016_2_20_30032

[50] "Killing Advanced Threats in Their Tracks: An intelligent Approach to Attack Prevention", T. Sager, The SANS Institute White paper, Jul 2014, [Online]. Available: https://www.sans.org/reading-room/whitepapers/analyst/killing-advanced-threats-tracks-intelligent-approach-attack-prevention-35302

[51] D. McMillen, "indicators of Compromise. Finding the footprints that attackers leave behind when they breach your defenses". IBM Security, New York, USA, SEL03041-USEN-00, 2015, [Online]. Available: https://pcatt.org/techblog/wp-content/uploads/2015/10/IndicatorsOfCompromise.pdf

[52] "Patterns of Compromise & Intelligence-Driven Threat Detection", D. Shackleford, Carbon Black Whitepaper, 2016, [Online]. Available: https://www.satisnet.co.uk/sites/default/files/cb_wp_patterns_of_compromise-threat%20intel.pdf

[53] M. Cloppert, 'SANS Digital Forensics and Incident Response Blog'. *Security Intelligence: Attacking the Cyber Kill Chain.* 2009 [Online]. Available: https://digital-forensics.sans.org/blog/2009/10/14/security-intelligence-attacking-the-kill-chain

[54] D. Bianco, 'Enterprise Detection & Response'. *The Pyramid of Pain.* 2013 [Online]. Available: http://detect-respond.blogspot.com.ee/2013/03/the-pyramid-of-pain.html

[55] "Tools and Standards for Cyber Threat Intelligence Projects", G. Garnham, The SANS Institute White paper, Oct 2013, [Online]. Available:

https://www.sans.org/reading-room/whitepapers/warfare/tools-standards-cyber-threat-intelligence-projects-34375

[56] "Detect, SHARE, Protect. Solutions for Improving Threat Data Exchange among CERTs", ENISA White paper, Oct 2013, [Online]. Available: https://www.enisa.europa.eu/publications/detect-share-protect-solutions-for-improving-threat-data-exchange-among-certs/at_download/fullReport

[57] "Standards and tools for exchange and processing of actionable information", ENISA White paper, Jan 2015, [Online]. Available: https://www.enisa.europa.eu/publications/standards-and-tools-for-exchange-and-processing-of-actionable-information

[58] D. Wilson. (2013, Sep. 17). The History of OpenIOC [Online]. Available: https://www.fireeye.com/blog/threat-research/2013/09/history-openioc.html

[59] "Using IOC (Indicators of Compromise) in Malware Forensics", H.-Y. Lock, The SANS Institute White paper, Feb 2013, [Online]. Available: https://www.sans.org/reading-room/whitepapers/forensics/ioc-indicators-compromise-malware-forensics-34200

[60] "Introduction to STIX", Oasis CTI TC Github repository [Online]. Available: https://oasis-open.github.io/cti-documentation/stix/intro

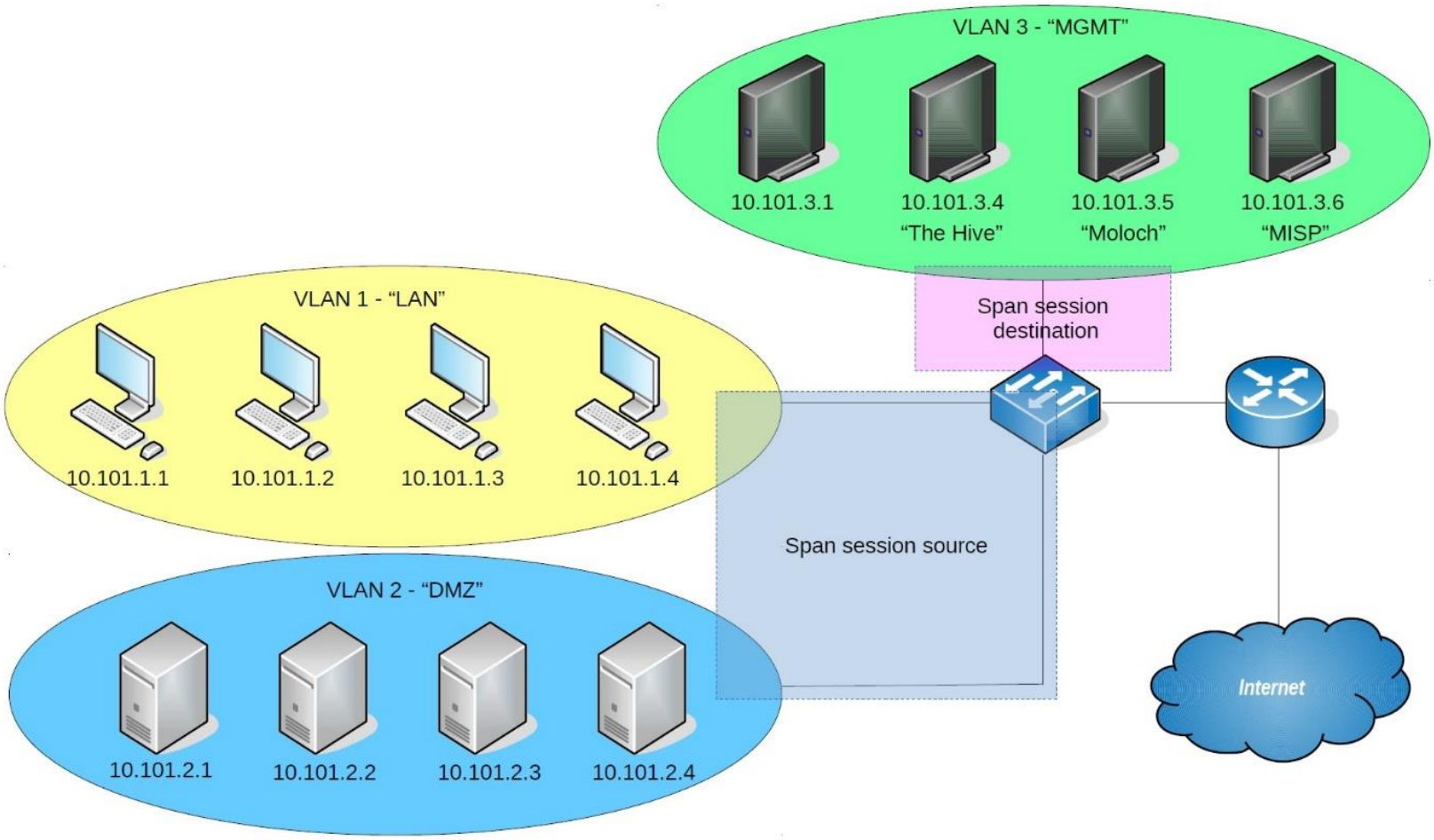
[61] "About MAEC", MAEC project on Github [Online]. Available: http://maecproject.github.io/about-maec/

[62] Anomali, Inc. (2017). *"What are STIX/TAXII"* [Online]. Available: https://anomali.cdn.rackfoundry.net/files/Anomali-STIX-TAXII.pdf

[63] K. Dunham. (2017, Jul. 6). Indicators of Compromise (IoCs) are Not Intelligence [Online]. Available: https://www.optiv.com/blog/indicators-of-compromise-iocs-are-not-intelligence

[64] National Council of ISACs. Member ISACs [Online]. Available: https://www.nationalisacs.org/member-isacs

Appendix 1 – Test environment description.



Appendix 2 – A Bash script written by author for partial automation of incident triage process .

```
#!/bin/bash

OIFS="$IFS"
IFS=$'\n'

#Security Incident Response Platform TheHive can be configured to automatically import malware-related information from the MISP events and store the obtained data as "Alerts".
#Alerts are stored in JSON format and can be accessed via TheHive API (with the help of curl command).
#This script is designed to work with TheHive alerts and use them to perform IOC hunt within given environment.

#The script works according to the following algorithm:
#It checks if TheHive has received any "new" alerts from MISP. If it has, the script:
# 1. Parses information contained in alert description to extract available threat-related tags and IOCs.
# 2. Searches network traffic for extracted IOC (currently limited to IP addresses / domain names) matches.
# 3. In cases when matches are found - determines importance of oragnization's hosts that can be potentially affected by a given threat.
# The script does this by comparing IP address of potentially affected host with network topology information, which is stored locally within a file named "zones".
```

```
# 4. Uses threat-related tags to estimates threat priority by referring to a local file named "escirt.taxonomy".
# The file contains results of simplified threat assessment process (different incident classes are assigned numeric values that describe corresponding threat priority).
# 5. Calculates potential incident priority in accordance with the following formula: Incident priority = threat priority * host importance.
# 6. Creates multiple cases describing potential incidents in TheHive:
# Cases describing potential incidents related to individual hosts that could be affected by corresponding threats; and
# "Parent" cases, which include the list of organization's hosts that could be potentially affected by given threat, and summarized impact of a particular threat that has been described in MISP.

#First, the script pulls data from TheHive alerts, which are marked with the "New" status and stores the data in variable "alerts".
alerts=$(curl -sH 'Authorization: Bearer API key' http://127.0.0.1:9000/api/alert | jq '.[] | select(.status == "New")')

#Each alert stored in TheHive has its own unique ID. The script will use alert IDs to analyze the data separately for every alert.
#The script parses data stored in variable "alerts", selects event IDs and passes their values to a variable "alertID".
alertID=$(echo $alerts | jq '.id' | tr -d \")

#Alerts contain various data provided by MISP. This script focuses on extracting malware-specific IOCs and taxonomy tags.
```

```
#The script examines each alert by its unique ID.
for i in  $alertID
  do

#The script queries TheHive to save information about a specific alert into variable "event".
    event=$(curl -sH 'Authorization: Bearer API key' http://127.0.0.1:9000/api/alert/$i)

#Next, the script parses data saved in variable "event" and extracts several pieces of information for different purposes.
#It begins with the name of corresponding MISP event. These names will be used to describe TheHive cases, when they will be created.
#The MISP event name is assigned to the variable "name".
    name=$(echo $event | jq '.' | grep title | cut -d'"' -f4)

#Possible incident severity depends on the class of the described threat (e.g. ransomware, trojan, etc).
#Such information can be obtained from tags, which are added to MISP events before they are published.
#Parse alert-related information and search for a tag that describes the class of threat in accordance with eCSIRT taxonomy (applied in MISP).
#Upon detection assign its value to variable "IOCclass"
    IOCclass=$(echo $event | jq '.tags' | grep csirt | tr -d \" | tr -d ' ')

#The script edits the extracted tag to make it human-readable.
    class=$(echo $IOCclass | cut -d'\' -f2)
```

```
#To support incident severity estimation, organization should conduct threat assessment on a regualr basis.
#A simplified example of threat assessment results have been saved into local file "ecsirt.taxonomy", where each type of potential threat is assigned a numeric value.
#To estimate severity of threat described in the alert, the script searches "ecsirt.taxonomy" file for matches with the value stored in variable "IOCclass", and picks a corresponding number.
#The number represents priority of a given threat according to threat assessment. Script passes number's value to variable "ThreatPriority".
    ThreatPriority=$(grep -F $IOCclass ecsirt.taxonomy | cut -d ',' -f 2)

#During previous steps various alert-related data has been passed to variable "$event" in json format.
#One of the passed components is a json object "artifacts", which includes different IOCs, published in MISP.
#To obtain alert-specific IOCs, the script creates a file "artifacts.alertID", extracts the "artifacts" object from the "$event" variable and saves it into the created file.
    echo $event | jq '.artifacts' | jq -c '.[]' > artifacts.$i

#IOCs can be used to determine organization's hosts that could have possible been affected by the examined threat.
#This script does this by analyzing network traffic (full packet capture), which is stored in elasticsearch database.
#To access elasticsearch database script will make use of Moloch API and compile different queries depending on the type of a particular IOC (either domain name, or IP address).
```

```
#The initial step in compiling a Moloch query is to determine the type of IOCs, stored in file "artifacts.alertID", and after that - its value.
    cat artifacts.$i | while read line
      do
#The script creates a varible "dataType" that will identify the type of IOC.
#The variable is assigned a corresponding value, which is extracted from every line of file that describes threat-related artifacts (e.g. "ip"/"domain")
        dataType=$(echo $line | cut -d: -f2 | cut -d, -f1)

#Next, the script defines a variable "data", which will store the actual value of a specific IOC (e.g. "192.168.10.10", or "www.maliciousdomain.bad")
        data=$(echo $line | cut -d'"' -f24 | cut -d'"' -f1)

#Depending on the determined IOC type, the script compiles a corresponding query and sends it to Moloch to search network traffic for IOC matches.
        if [ $dataType = '"ip"' ]; then

#If matches are detected, the Moloch returns an answer in JSON format, describing relevant network connections.
#This allows script to filter out IP addresses of hosts belonging to organization's network range and store the list of such addresses in variable "queryIP".
          queryIP=$(curl --digest -u username:password "http://10.101.3.105:8005/connections.json?expression=ip.dst%20%3D%3D%20${data}&date=-1" | jq '.nodes' | grep id | grep -v ${data} | cut -d'"' -f4)
```

```
#The script saves determined IP adresses into local file named "affected.addresses.alertID".
#This is required, because this list should be slightly modified.

            for address in $queryIP
              do
                echo $address >> affected.addresses.$i
              done

        elif [ $dataType = '"domain"' ]; then
         querydomain=$(curl   --digest   -u   username:password "http://10.101.3.105:8005/connections.json?expression=http.uri%20%3D%3D%20${data}&date=-1" | jq '.nodes[] | select(.type == 1)' | grep id | cut -d'"' –f4)

            for address in $querydomain
              do
                echo $address >> affected.addresses.$i
              done
        else
          echo bad match
        fi
      done

#Different queries may return the same result and thus create duplicated lines in file "affected.addresses.alertID".
#After the file "affected.addresses.alertID" is populated with data provided as results of both executed queries (IP/domain), the script sorts its content and keeps only unique IP addresses in variable "affected.hosts.alertID"
```

```
#Such list will be used during creation of "parent" cases for possible incidents.
  sort affected.addresses.$i | uniq > affected.hosts.$i

#Additionally, the script calculates the amount of unique hosts that are potentially affected by the specific threat, and keeps this number in variable "number"
    number=$(wc -l affected.hosts.$i | cut -d' ' -f1)

#"Host importance" is one of the factors that affect severity of a particular incident.
#Once IP addresses are determined, the script can estimate their importance.
#Importance of potentially affected hosts can be established in accordance with network segment that they belong to.
#To determine host importance the script uses hosts's IP address and file "zones" that is stored locally.
#File "zones" includes IP address ranges of network segments and numeric values, which describe importance of different network segments.

#To perform comparison of host and network segment's IP addresses, they will have to be translated into decimal integers.
#For this purpose the script uses 2 functions (published by Dennis Williamson on www.stackoverflow.com):
# 1.The script defines a function "ip2dec" to translate IP addresses into decimal integers and check if IP address fits into a defined range.
    ip2dec () {
    local a b c d ip=$@
    IFS=. read -r a b c d <<< "$ip"
```

```
    printf '%d\n' "$((a * 256 ** 3 + b * 256 ** 2 + c * 256 + d))"
    }

# 2.The script also defines function "dec2ip" for opposite purpose: translating decimal integers to IP addresses.
    dec2ip () {
    local ip dec=$@
    for e in {3..0}
    do
        ((octet = dec / (256 ** e) ))
        ((dec -= octet * 256 ** e))
        ip+=$delim$octet
        delim=.
    done
    printf '%s\n' "$ip"

#The script translates IP addresses of potentially affected hosts into integers and assign integer values to variable "target"
    cat affected.hosts.$i | while read line
do
        target=$(ip2dec "$line")

#The script translates lower and higher IP addresses of network ranges stored in file "zones" for comparison purposes.
        cat zones | while read line
        do
          min=$(echo $line | cut -d- -f1 | cut -d' ' -f2)
          max=$(echo $line | cut -d- -f2 | cut -d, -f1)
          mindec=$(ip2dec "$min")
          maxdec=$(ip2dec "$max")
```

```
#The script extracts numeric values describing importance of network segment (from the "zones" file) and assign its value to variable "HostImportance"
          HostImportance=$(echo $line | cut -d, -f2)

#From now on, the script can use available information to estimate possible incident priority.
#The script does it by multiplying "Threat Priority" value with "Host Importance" value and saving the result into variable "IncidentPriority".
          IncidentPriority=$((ThreatPriority*HostImportance))

#IP addresses of potentially affected hosts are confronted against network segment address ranges.
#If match is detected, the host is assigned an importance value and possible incident priority is calculated.
          if [ $target -gt $mindec ] && [ $target -lt $maxdec ]; then
            zone=$(echo $line | cut -d' ' -f1 )
            ip=$(dec2ip "$target")

#Additionally, the script creates a file "impact.alertID" and begins filling it with alert-related incident priority values, calculated for each unique potentially affected host.
#This file will be used during creation of corresponding "parent" TheHive case.
          echo ${IncidentPriority} >> impact.$i
```

```
#When data related to potentially affected host importance is calculated, the script creates individual cases in TheHive for each unique host that can be potentially affected.

        curl -XPOST -u username:password -H 'Content-Type: application/json' http://127.0.0.1:9000/api/case -d '{
          "title": "[MISP] '$name', incident priority ='${IncidentPriority}' ",
          "description": "This case is automatically generated by script that searched network traffic for IOCs obtained from MISP event '$name'. \n\r Incident is detected in '$zone' network segment. \n\$
          "severity": 3,
          "tlp": 3,
          "tags": ["eCSIRT:'$class'", "script_generated"]
        }'
          fi
        done
      done

#After the script has created cases for each determined individual host, it calculates the potential impact of each MISP event / new alert in TheHive.
#The script does this by summarizing priority values of possible incidents related to individual hosts and stores the result in a file "impact".
    impact=$(numsum impact.$i)

#After the script has created cases describing possible incidents related to individual hosts, it creates a "parent" case.
```

```
#The "parent" case includes the list of potentially affected hosts and summarized impact of a particular threat, that has been described in MISP event / TheHive alert.
#To send the list of affected hosts to TheHive using API, the list has to be converted to a single string.
#A variable "hosts" is used by script to store the value of such string.
    hosts=$(cat affected.hosts.$i | tr '\n' ';')

#The script uses TheHive API to create a "parent" case for a specific MISP event / TheHive alarm.
    curl -XPOST -u username:password -H 'Content-Type: application/json' http://127.0.0.1:9000/api/case -d '{
      "title": "[MISP: Parent case] '$name', estimated impact = '$impact'",
      "description": "This case is automatically generated by script that searched network traffic for IOCs obtained from MISP event '$name'. \n\r Summarized priority of possible incidents related to thi$
      "severity": 3,
      "tlp": 3,
      "tags": ["eCSIRT:'$class'", "script_generated"]
    }'
#In the end the script removes files that have been temporarily saved by him on a local host.
rm artifacts.*
rm affected.addresses.*
rm affected.hosts.*
rm impact.*
rm hive.input.*
  done
```