UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE INFORMÁTICA

Assessment of cyber threats discovered by OSINT

Francisco Contreras Leão Gomes

Mestrado em Segurança Informática

Dissertação orientada por:
Prof.ª Doutora Ana Luísa do Carmo Correia Respício
Prof. Doutor Pedro Miguel Frazão Fernandes Ferreira

2021

Acknowledgments

I want to start by expressing my sincere gratitude to my advisors, Prof. Doutor Ana Respício and Prof. Doutor Pedro Ferreira, for always pushing me to keep working, for helping me and advising through the whole process. I would also to thanks LASIGE and its researchers, Fernando Alves and Nuno Dionísio, for the knowledge sharing and for working with me in the integration phase, this was not possible without them.

To BNP Paribas Securities Services Portugal and my coworkers, for supporting and allowing me to keep studying and working, thank you.

And, finally, a special thanks to my family, for the unconditional support, for motivating me during this delicate phase and for understanding by absence in some special events.

To my parents and grandparents.
To my family.

Resumo

Apesar dos altos níveis de maturidade das ferramentas, técnicas e procedimentos de inteligência de ciberameaças, comumente denominado por CTI (Cyber Threat Intelligence, na literatura anglo-saxónica), ainda existem lacunas que devem ser consideradas e abordadas. Mais de 50 % da população mundial está online e com tendência a crescer, ao mesmo tempo que a pandemia COVID-19 impulsiona a adoção em larga escala de tecnologia nas mais diversas áreas. Este contexto, aliado às tecnologias emergentes (por exemplo: computação na nuvem, IoT, 5G), viabiliza e permite ciberataques mais complexos e rápidos. A segurança ainda não é um requisito abordado nas fases iniciais de projeto, já que os produtos precisam de ser rapidamente colocados no mercado, deixando alvos vulneráveis no ecossistema da Internet. Estima-se que o cibercrime provoque danos de 6 biliões de dólares em 2021, crescendo 15 % ao ano, posicionando-se como a terceira maior economia do mundo, atingindo 10.5 biliões de dólares em 2025 [1]. Os ciberataques a infraestruturas críticas foram considerados o quinto maior risco em 2020, já que as grandes indústrias e setores estruturais são alvos aliciantes. Por outro lado, a probabilidade de deteção e acusação é estimada em 0,05 % nos EUA [2]. Para combater esta ameaça e reduzir o risco, é essencial que todas os participantes no fluxo de CTI se unam para melhorar a coordenação e a cooperação, para reduzir o tempo entre a geração de CTI e sua disseminação, para alcançar o equilíbrio entre a disseminação no tempo e a alta qualidade dos dados de CTI. A insuficiente qualidade dos dados de CTI é uma grande barreira para atingir estes objetivos. A maioria das plataformas consome dados de fontes pagas ou OSINT, recolhendo, filtrando, analisando e agregando, geralmente com pouca ou nenhuma avaliação da qualidade dos dados, aumentando assim a pressão sobre os analistas de cibersegurança, que lidam diariamente com uma infinidade de alertas. Os IoCs devem por isso passar por um processo de avaliação e serem pontuados, para que os consumidores de CTI possam agir mais agilmente e em conformidade. De acordo com a pesquisa ENISA 2020 CTI [3], apenas 4 % das plataformas de CTI conseguem implementar processos para medir a eficiência de CTI.

Tendo como ponto de partida uma componente desenvolvida no âmbito de um projeto europeu H2020 (DiSIEM OTD [4]), desenvolvida na unidade de investigação *LASIGE*, do Departamento de Informática da Faculdade de Ciências da Universidade de Lisboa, iniciámos a pesquisa de trabalho relacionado, bem como de métricas que nos ajudas-

sem a classificar dados, mais concretamente *BigData*. Esta ferramenta [4], apesar de não oferecer qualquer avaliação dos dados ao utilizador, é responsável pela recolha, filtragem, classificação e agregação de dados proveniente da plataforma social *Twitter*, sendo esta sua a única fonte *OSINT*(Open Source Intelligence). Outros trabalhos semelhantes propõem a introdução de algumas propriedades para auxiliar a medição da qualidade dos dados, apesar de não o fazerem em concreto. No contexto da qualidade dos dados, a norma ISO/IEC25012:2008 [5] oferece um modelo para avaliar dados estruturados, onde quinze características são propostas de acordo com dois pontos de vista: 1) qualidade dos dados; 2) capacidade de reter os dados num sistema, havendo características que são abrangidas por ambos os pontos. O primeiro ponto contém diversas características que são relevantes para este trabalho e que nos serviram de inspiração. Foi igualmente importante o estudo dos diferentes conceitos associados à CTI e aos IoCs, para melhor entender como podíamos explorar a integração das métricas de qualidade dos dados nas plataformas e formatos existentes, tanto comerciais, como comunitárias.

A nossa proposta contém doze métricas para avaliar a qualidade dos dados OSINT e representá-la plataforma DiSIEM OTD [4], subdivididas em dois grupos: 1) sete métricas que avaliam as contas da plataforma *Twitter* em diversas vertentes, 2) cinco métricas responsáveis por avaliar os dados de cada *tweet* ou agrupamento de *tweets*. Para o demonstrarmos e avaliarmos experimentalmente, desenvolvemos uma aplicação arquitetada em três fases: fase 1) DiSIEM OTD [4]: definição da infraestrutura a monitorizar e respetivas keywords; fase 2) Utilizando os dados extraídos na fase 1, o MCD (Multitask Cyberthreat Detection) [6] [7] efetua a normalização de cada tweet (removendo carateres especiais), identifica entidades como organizações/versões de aplicações/ameaças/identificadores de vulnerabilidades, e classifica a relevância de cada *tweet*, estimando a probabilidade de conter informação valiosa sobre um ativo de interesse; fase 3) Utilizando os dados processados na fase 2, a nossa aplicação consulta a *API* do *Twitter* para enriquecer os dados com propriedades e métricas públicas dos *tweets* e das contas do *Twitter*, calculando todas as doze métricas, posteriormente reproduzidas em três ficheiros *CSV* distintos, estando igualmente preparada para exportar ficheiros nos formatos *MISP* e *STIX*, abrindo assim a porta a integrações com outras plataformas. Mesmo existindo certas limitações na utilização da *API* do *Twitter*, o conjunto de dados não foi suficientemente largo para haver impacto na aplicação. Apesar destas fases estarem intrinsecamente relacionadas, o facto de não terem existido desenvolvimentos na plataforma DiSIEM OTD [4] impossibilitou a integração das métricas na mesma. Recorremos, por isso, a uma plataforma externa para desenhar as maquetes e realizar a prova de conceito, tendo sempre como base a DiSIEM OTD [4]. Durante a avaliação experimental das métricas propostas, utilizámos um conjunto de dados recolhidos pela DiSIEM OTD [4] no período de 15/Julho/2021 a 14/Setembro/2021. As entrevistas conduzidas com especialistas na área da cibersegurança revelaram a utilidade das métricas e da sua apresentação aos utilizadores.

No decorrer deste projeto, surgiram diversas ideias que gostaríamos de executar posteriormente e que beneficiariam em grande medida a plataforma que serviu de base à dissertação e toda a comunidade da cibersegurança.
Em suma, esta dissertação apresenta uma visão geral das metodologias e tecnologias de CTI existentes, propondo uma solução a ser adotada e integrada em ferramentas de CTI para avaliar, qualificar, pontuar e aconselhar os analistas de cibersegurança, que podem ser utilizadas para construir relatórios, ou integradas em TIPs (Threat Intelligence Platform), agilizando assim o tratamento do risco e a resposta a potenciais incidentes.

Palavras-chave: cibersegurança; informações de ciberameaças; informações de fonte aberta (OSINT); indicadores de comprometimento (IoC); qualidade dos dados;

Abstract

Despite the high maturity levels of CTI (Cyber Threat Intelligence) tools, techniques, procedures and frameworks, there are still gaps that must be considered and addressed. More than 50% of the world's population is now online and growing, as the COVID-19 pandemic is pushing the large-scale adoption of technology in the most diverse areas. This context, aligned to the emerging technologies (e.g.: Cloud-computing, IoT, 5G) is enabling, allowing, and amplifying more complex and faster cyber-attacks. "Security-by-design" is not yet the main principle, as products need to be quickly deployed into the market, delivering vulnerable targets into the Internet ecosystem. It is estimated that cybercrime inflict damages of 6 billion USD in 2021, growing 15% per year, positioning it as the world' third-largest economy, reaching 10.5 billion USD in 2025 [1]. Cyberattacks on critical infrastructures was considered the fifth top risk in 2020, as structural industries and sectors are juicy targets. On the other hand, the likelihood of detection and prosecution is estimated to be 0.05% in the USA [2]. To fight this threat and reduce the risk, it is essential that CTI parties join forces to improve coordination and cooperation, to reduce the time between the generation of CTI and its dissemination and achieve the balance between CTI in-time-dissemination and high-quality CTI. The quality of CTI is a huge barrier: most of the platforms ingest data from paid feeds and OSINT sources, gathering, filtering, analyzing, and aggregating, usually with little or no data-quality assessment. This increases the pressure on cyber-security analysts, who deal with plenty of generated alerts. IoCs (Indicator of Compromise) must go through an assessment process and be scored, so CTI consumers can decide and suit the measures accordingly. According to ENISA 2020 CTI survey [3], only 4% of CTI users can implement processes to measure CTI efficiency. This dissertation presents an overview of the existing CTI methodologies and technologies, proposing one solution to be adopted and integrated in CTI tools to assess, qualify, score and advise cyber-security analysts.

Keywords: cybersecurity; cyber threat intelligence (CTI); open source intelligence (OSINT); indicators of compromise (IoC); data quality;

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In an era where OSINT is more and more widely available, it is essential to assess data quality and use the correct information to take good decisions and act correctly. Big Data and Cyber Security found themselves united to improve cyber risk analysis and management.

1.1 Motivation

Cyber risk management has become a critical activity in day-to-day operations of organizations, independently of the offered products/services. Despite the tactics, guidelines and well established standards for cyber security risk management, the assessment of cyber risk remains an activity that presents great challenges, essentially due to the complexity of estimating the uncertainty of emerging cyber threats. This need boosted the development of CTI, that can be described as knowledge based on evidences on threats to information technology assets. In particular, the OSINT intelligence, combined with internal information about the IT infrastructure, provides knowledge to support the implementation of defense controls and response to cyber incidents. However, using the information from external sources to make informed decisions requires selecting which is the most relevant, accurate and reliable information.

According to the SANS CTI Survey 2021 [8], 25.9% of the respondents are not satisfied with the ”Relevance of threat data and information” and 32.9% are not satisfied with the ”Cleanliness and quality of data” (see Figure 1.1).

	Very Satisfied	Satisfied	Total Satisfied	Not Satisfied
Visibility into threats and IOCs	14.0%	56.1%	70.2%	24.1%
Searching and reporting	14.9%	52.6%	67.5%	23.2%
Reports (strategic and operational level)	15.4%	51.8%	67.1%	24.1%
Timeliness of threat data and intelligence	14.9%	51.8%	66.7%	29.4%
Relevance of threat data and information	13.6%	52.6%	66.2%	25.9%
Automation and integration of CTI information with detection and response systems	16.2%	49.0%	65.2%	28.1%
Context	12.7%	46.9%	59.6%	32.9%
Cleanliness and quality of data	10.1%	49.6%	59.6%	32.9%
Integrated data feeds	12.3%	44.3%	56.6%	30.3%
Comprehensiveness of coverage	14.5%	41.2%	55.7%	36.8%
Analytics	13.2%	39.5%	52.6%	31.6%
Location-based visibility	7.9%	37.3%	45.2%	32.9%
Identification and removal of expired IOCs and other old data	8.8%	36.0%	44.7%	45.6%
Machine learning	9.2%	23.2%	32.5%	37.7%
Other	2.2%	8.3%	10.5%	3.1%

Figure 1.1: CTI effectiveness - SANS CTI Survey 2021 [8]

1.2 Objectives

This project aims to identify metrics that allow assessing the relevance and reliability of information extracted from OSINT. More specifically, it is intended to assess the severity of threats discovered through OSINT for a risk assessment, associating the specificity of a factual infrastructure. The implementation was based on DiSIEM OTD[4], that by monitoring *Twitter*, is responsible for OSINT collection and data clustering, and where the proof-of-concept took place.

1.3 Contributions

This dissertation identifies 12 metrics to assess the trustworthiness of discovered threats and proposes two new ones to assess *Twitter* accounts and IoCs, helping analysts to prioritize the response, i.e., the risk treatment and response to potential incidents. Our proposal was evaluated through interviews with two cybersecurity specialists, who considered the proposed metrics and the approach very useful to implement in a platform like DiSIEM OTD [4], besides the potential contributions and integrations that it can trigger, as documented in Section 7.1.

1.4 Structure of the document

This document is organized as follows:

- Chapter 2 provides an overview of the related and background work, regarding this thematic.
- Chapter 3 explains some of the important concepts in this field.
- Chapter 4 introduces the new proposals of possible dimensions to measure the data quality.
- Chapter 5 develops the proposed architecture and features.
- Chapter 6 details the experimental evaluation, the obtained results and analysis.
- Chapter 7 presents the concluding remarks, some limitations of our work and improvements for future work than can be done based on the obtained results.

Chapter 2

Related work

In this chapter, we present some background and review some important literature that is relevant to the assessment of cyber threats, mainly related with OSINT sources and its analysis, with the objective of supplying the whole IoC chain with enriched and fast-analysis information.

2.1 Twitter

Twitter is a social network with more than 192 million daily active users, as of Q4 2020 [22]. More than a personal platform for content sharing and promotion, it is widely used by hackers, vendors, cybersecurity analysts and enthusiasts to discuss and disclose exploits, vulnerabilities, threats and attacks. The early detection is very important, as it allows prioritizing the response [23], i.e., the risk treatment and response to potential incidents.
On SYNAPSE [9], Alves et al. managed to build a pipeline that collects, filters, preprocesses, extracts, classifies and clusters the information, producing IoCs to be consumed by analysts and its cyber threat management platforms, represented in Figure 2.1. Despite the classification module that measures the relevance of each tweet, this end-to-end system doesn't identify the metrics that allow the evaluation of the relevance or trustworthiness of the information extracted from Twitter. Also, as the IoCs are not enriched with this kind of information, analysts are not aware of the classification score and cannot correctly appraise the threat's severity and the associated risk. The system works with pre-selected Twitter accounts and keywords, also not (re)evaluated in time.

Also focused on Twitter, Mittal et al. [24] developed CyberTwitter, a system to gather and analyze cybersecurity intelligence on Twitter and serve as a OSINT (Open–source intelligence) source. The system is capable of identifying, tag and extract real world conceptual entities related to cybersecurity vulnerabilities such as means of an attack, consequences of an attack, affected software, hardware, vendors, etc. using a Security Vul-

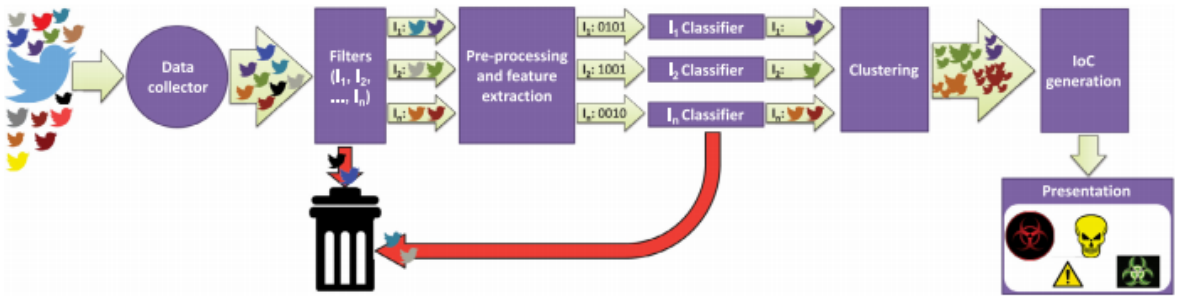


Figure 2.1: SYNAPSE pipeline architecture, extracted from [9]

nerability Concept Extractor (SVCE). Intelligence from these terms and concepts is then converted to RDF [25] statements using intelligence ontology like UCO[26] or CASE[27], that is then related to the user system profile (USP, i.e. where the infrastructure software and hardware (assets) can be defined). At this moment, *'Intelligence'* has an actionable information for the human analyst, which makes them aware about a new threat or vulnerability in a software / hardware that they list in their user system profile, without temporal dimension. To mitigate that, time dimension was introduced, by adding the following properties at each term:

- **counter**: number of tweets collected with a given CTI (i.e.: number of tweets in a cluster).
- **beginTime**: timestamp of the first known CTI of a given cluster.
- **lastTime**: timestamp of the last known CTI of a given cluster.
- **hasVulnerability**: Boolean – if the term refers a vulnerability or not.
- **productInUSP**: if the term refers to an asset defined in the user system profile.
- **isCurrentlyValid**: 'True' if the intelligence term is 'valid and current'. A valid and current intelligence is a one that gives details about an open, temporally significant vulnerability or threat in an affected asset.

Here, valid and current CTI is associated with sets of rules and can be consumed by platforms or can trigger alerts to users. Even if this project doesn't classify CTI, the timeliness/novelty dimension, associated with the tagging of the means of an attack, consequences of an attack, affected software/hardware/vendors, it could be a great start to evaluate CTI quality from OSINT sources. In addition to that, it would be great to give weight to CTI source score, that will vary each time that one CTI is consumed from that source. Users may also be able to customize the weight of each dimension.

The assessment of Twitter accounts is the target of Tundis et al. [28] work, where an approach for the automated assessment of the OSINT sources is proposed, acting themselves as an additional criterion for the relevance of CTI. Two feature sets that characterize the OSINT source have been defined and an experimentation was conducted by training 5 regression models on both feature sets to predict the relevance score for OSINT sources. The 1st feature set, centered on meta-data, has been selected by considering 3 aspects:

1. **Profile related features:** characteristics of a Twitter profile that are directly associated with the user profile (e.g. registration date, the user's specified location, number of followers and so on);

2. **Social graph related features:** relations between them within a group or community of connected profiles (followed/following, retweets and mentioned/mentions);

3. **Tweet related features:** other features, which are specifically associated to a single tweet, that provide additional information on the user's behavior with regards to the published Tweets. The 2nd feature set is based on the word embedding technique *"doc2vec"* algorithm, that strives when determining the similarity between different textual data.

A scoring function to quantify the relevance of an OSINT source with regards to CTI in particular consideration of the timeliness is proposed. For calculating the score for a single CTI term *i*:

$$r_i \quad = \quad score(t_i) \quad = \quad \begin{cases} 1 - 0.5 \cdot \left(\frac{t}{C-1}\right)^2 & s \cdot (c-1)^{1.25} \leq t \leq s \cdot c^{1.25}, \\ & 0 < c \leq C, c \in N \\ 0.5 \cdot 0.5^{\frac{t}{s}} & s \cdot C < t \end{cases}$$

, where t is the time difference in seconds between the first time a CTI term has been observed and the moment it is mentioned again by other source, C = 5 and s = 86400 (seconds in a full day) - both empirically determined.

All terms are then aggregated per source *I* in order to get a single score to each source:

$$cti\_relevance\_score(R_I) = \frac{1}{|R_I|} \sum_{i=1}^{|R_I|} r_i \cdot \frac{log(|R_I|)}{log(|R|)}$$

, where *R* represents the full set of all scores and *RI* the scores for intelligence shared by source *I*.

After all sources have been scored with a *"CTI Relevance"*, they are used to train a model to predict the relevance. The results indicate that the *"CTI Relevance Score"* can be predicted from CTI source features using the presented regression models. Such a score can be used to increase the timeliness of alert generation, an immediate alert can be generated instead of waiting for a second occurrence of that intelligence from a different source.

2.2 Data quality

Focusing on the data quality assessment, Schlette et al. [29] propose to extend the IoC STIX format and include some data quality dimensions, based on the evaluation of the STIX objects collected from feeds. Adjustable aggregation parameters enable users to define the weight of each data quality dimension, when calculating the global score of each STIX object. The weighted average, where each dimensional score is weighted, can be adjusted by each platform consumer, ensuring that the quality scores represent their individual preference of the dimension's importance.

Azevedo et al. [30] propose PURE, an automated solution for enrichment and quality IoC creation from OSINT. Their work focuses mainly on reducing the quantity of information that reaches a security analyst, increasing the quality of intelligence that arrives to an analyst and facilitating the automation of the generation of improved intelligence. This is done by:

1. collecting threat intelligence from different OSINT sources
2. normalizing it to a single IoC format (MISP, STIX, etc.)
3. removing duplicates and IoCs that do not provide new information
4. correlating and aggregating (correlates different IoCs and generates new enriched ones).

However, no data quality assessment is applied during that process, even if four(4) main data quality dimensions are proposed:

- **completeness**: how much the information contained in an intelligence artifact allows the identification of an attack.
- **accuracy**: how much the intelligence reduces the number of false positives.
- **relevance**: how much the intelligence relates to the specific purpose for which it is intended.
- **timeliness**: measures the time between the creation of an intelligence artifact and when it reaches its target, either human or defensive infrastructure.

Chapter 3

Concepts

3.1 Threat Intelligence

3.1.1 Definition

Threat Intelligence (TI) is the practice of acquiring, collecting, processing evidence-based knowledge, from multiple sources, about threats and its actors, besides their motives, intents and capabilities. The main purpose is to help organizations improve security and to agile the decision-making process. Advantage is secured over the adversaries, by having sufficient understanding for mitigating a harmful event. Having its roots in the military, where many intelligence frameworks and threat intelligence professionals started, CTI often uses military terms and techniques.

Figure 3.1: Intelligence, TI, CTI, extracted from [10]

The Cyberspace is like a military battlefield to a commander, where the weapons are virtual, the soldiers are behind the screen, but the consequences are real and can be catastrophic. The CTI added value does not simply apply to detecting and response, as Figure 3.2 demonstrates.

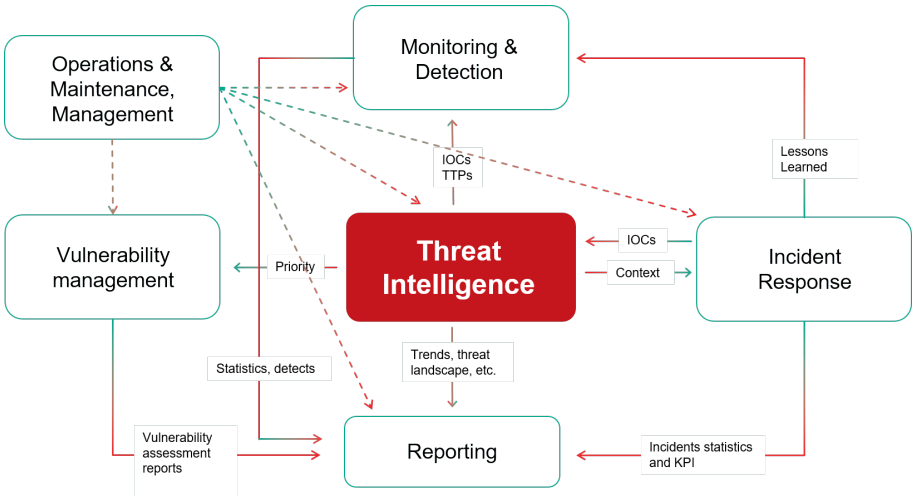


Figure 3.2: Threat intelligence–driven security operations, extracted from [11]

3.1.2 Information vs Intelligence

Raw and unfiltered information is not Intelligence by itself. To be considered Intelligence, data must be relevant, actionable, contextual and as timely and accurate as possible [31]:

- **relevant**: It must impact the organization in some way.
- **actionable**: Concrete steps can be taken by security teams to protect the organization.
- **contextual**: There should be enough evidence included to enable an intelligence analyst to effectively rank the threat.
- **timely**: It should be received with enough time to do something with it, otherwise is near useless.
- **accurate**: It should be reliable and detailed.

The transformation flow begins with the information collection, proceeds with the data processing/analysis and, with the help of evaluation, aggregation and interpretation tools, Information is transformed into Intelligence, that is then deployed and disseminated, ready to be consumed, as represented in Figure 3.3.

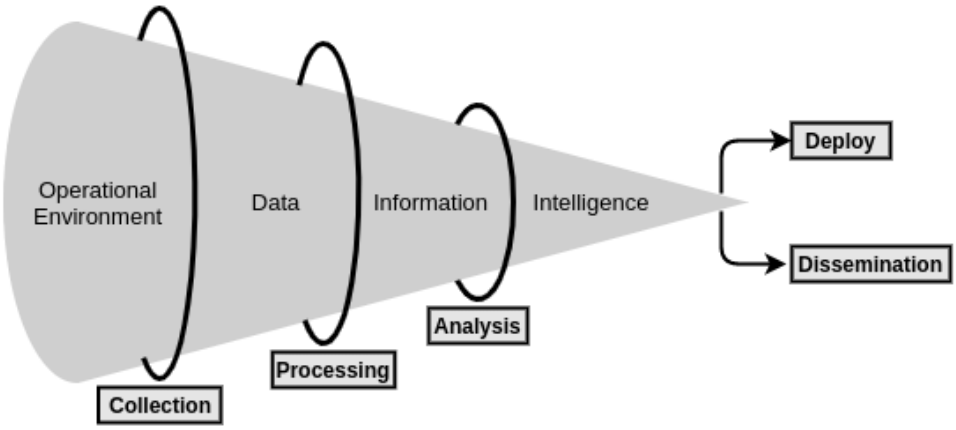


Figure 3.3: From Information to Intelligence, extracted from [12]

3.1.3 Importance

Figure 3.4 represents a pyramid with different threats and the common approaches to address it. Most companies and organizations are focusing their efforts on the base of the pyramid, by integrating CTI feeds into SIEMs, IPS/IDSs and firewalls, without taking full advantage of it. This happens because this type of data is cheap and they don't know what to do with all the remaining data, how to prioritize and what to ignore.

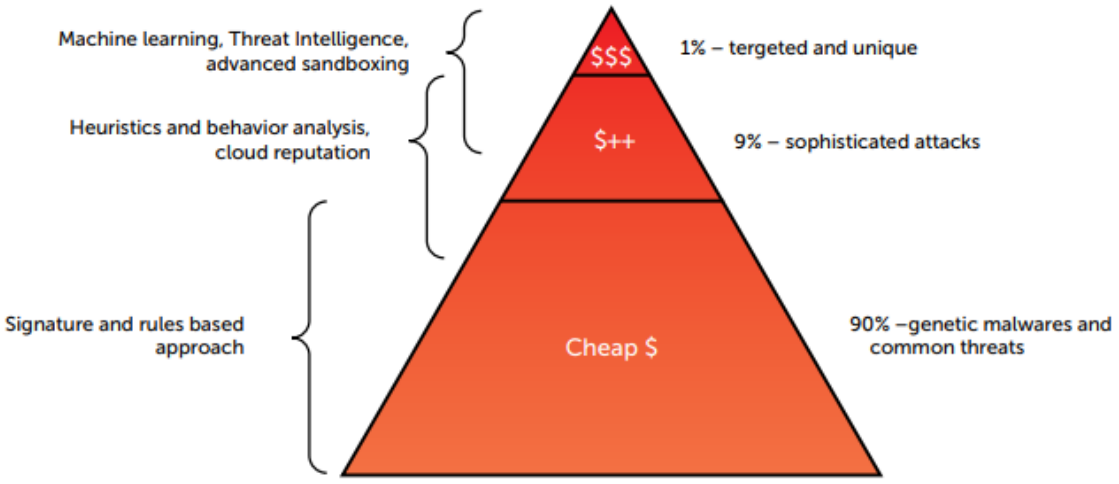


Figure 3.4: Addressing different threats with security technologies, extracted from [13]

But what enables companies/organizations to access threat intelligence from the bottom to the top of the pyramid represented in the Figure 3.4, is the sharing between entities, the exchange of intelligence and the contribution that every of each party can give and receive. As Cyber-Trust describes [10], all the parties can take advantage and benefit from it:

- Increased situational awareness: enhances security by leveraging the capabilities, knowledge and experience of partners.
- Improved security posture: it becomes easier for partners to identify threats, affected system and to implement both protection and recovery measures/protocols.
- Knowledge maturing: as CTI is shared and flows across parties, it is enriched with all the contributions, that no matter how small is, if the volume of participants is large, it ends up having the expected result.
- Increased defensive agility: reduce the time to act and the likelihood of a successful attack by continuously exchanging information about new vulnerabilities and, TTPs and its evade detection methods.

3.1.4 Lifecycle

The CTI lifecycle is the iterative and adaptable process through which Intelligence is obtained, produced, and made available to end-users. Most Intelligence is produced from raw information by way of the Intelligence cycle, that consists of five steps, illustrated in Figure 3.5:

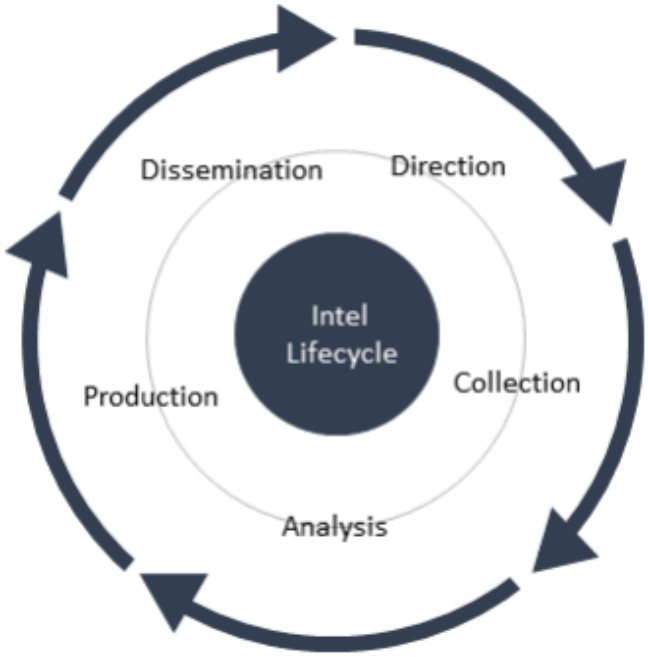


Figure 3.5: CTI lifecycle, extracted from [14]

This model, or a variant of it, is used by Military Intelligence [31] and it was adopted for the Cyberspace. The five steps flow around to help the organization succeed, by producing actionable Intelligence and therefore making the security team more effective.

- **Direction**: crucial planning stage as it sets the roadmap for the CTI operation. The team must define and set the sources, scope, requirements, methodology and goals and objectives of the Intelligence activities.
- **Collection**: gather data that fulfill the requirements. Normalize and enrich raw data, eliminate duplicated information, by:
  - Filtering: remove what does not add any value, i.e. filter out irrelevant, useless or worthless data.
  - Normalization: transform the collected data into uniform and supported formats.
  - Indexing: to enable search and navigation through large volumes of data in an efficient way.
  - Enrichment: add relevant and contextual metadata to the gathered data.
  - Prioritization: rank data to drive the analysis efforts.
- **Analysis**: evaluate, assess and interpret the information, in terms of confidence, relevance, likelihood, risk and impact, to get valuable recommendations for the stakeholders.
- **Production**: produce intelligence that fits in one of the levels discussed in Section 3.1.5.
- **Dissemination**: analysis translation into a digestible format and delivery and/or deploy the Intelligence across the platforms, stakeholders, feeds.

This ends in a feedback loop to encourage continuous improvement, where a report should be provided to determine whether adjustments need to be made.

3.1.5 Levels and Use Cases

Threat Intelligence can be broken down into three unique levels (see Figure 3.6), that deal with different objectives and uses, answering to different questions and being useful to distinct audiences. Figure 3.6 illustrates the intelligence levels across the pyramid, proposed in [15] :

- **Strategic**: focused on high-level trends and attacker's motives, it is used to report to top-level management and senior decision makers, to help them understanding the business risks and assist in the formulation of long-term strategies, as the name suggests. It is usually expressed in plain language or less technical and disseminated on a monthly/quarter basis in the form of reports.

Figure 3.6: CTI levels [15]

- **Tactical**: technical in nature, it outlines the TTPs, which are particularly useful for network operation teams, like SOCs, to understand how the organization might be attacked and how to defend/mitigate against. The main deliverable are IoCs, used on signature-based systems.
- **Operational**: it provides technical and specialized insights to help IR (Incident response) and VM (Vulnerability managament) teams to have knowledge about the nature, intent and timing of attacks, answering to ”when, where and how”.

Its value, use-cases and capability are different throughout the life-cycle [32], as CTI type varies, as illustrates Table 3.1.

CTI level	Data type	Answers to	Sources	Potential users	CTI value
Strategic	Intelligence	Who, Why	Intel. reports	Managers Management boards Committees CISO,CIO,CTO	- Understand adversary intent/ motivation - Calculate business risk - Determine targets, approaches, investments
Operational	Enriched data Information	When Where How	Intel. platforms	CSIRT Forensic Analysts IR/VM team	- Holistic understanding - Historical perspective - Threat hunting
Tactical	Data Filtered data	What	IoCs	NOC/SOC Firewall IDS/IPS	- Prioritize incidents - Asset patching - Monitoring

Table 3.1: CTI types

3.1.6 Sources of Intelligence

Beyond the CTI obtained from internal sensors and monitoring systems, it may origin from different sources:

- **HUMINT (Human Intelligence)**: based on a relationship between the intelligence agent and the human source.
- **CHIS (Covert Human Intelligence Sources)**: similar to HUMINT but when the relationship has been established with the strict purpose of gaining access to some restricted information.
- **SIGINT (Signals Intelligence)**: derived from the interception of signals, whether between people or between devices
- **OSINT (Open Source Intelligence)**: intelligence that is produced from OSINF, i.e: publicly available and unclassified information,like websites, blogs, discussions forums, social networks. Figure 3.7 represents the different sources and links across.

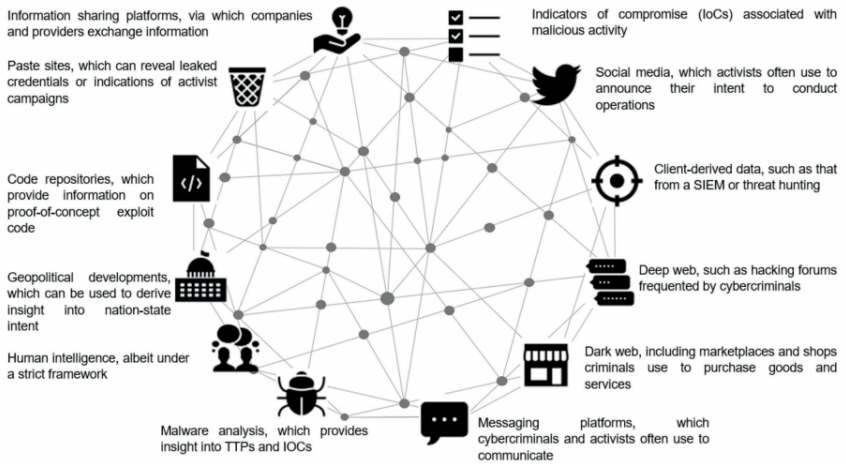


Figure 3.7: CTI OSINT sources [15]

  - **SOCMINT (Social Media Intelligence)**[33]: Social networks, like Twitter, allow users, cyber organizations and vendors to communicate and share vulnerabilities, exploits and anomalous activities. It is a powerful source of valuable information, that needs to be carefully treated to be effective.
  - **CTI Feeds (paid/free)**: organizations can develop their own ways of collecting and analyzing OSINT, but many trust and rely on feeds, either free or paid. Feeds are usually accessible with a simple HTTP or WS/API request and delivered by ONGs, national and organizational CSIRTS or vendors. They can contain IP/URL/hash/domain blacklists, tied to scanning, phishing campaigns, malware bots and other forms of threats.

3.1.7 Feeds

CTI feeds can provide low-level information such as blacklists (with hash values, IP addresses or domains) that can be helpful to block unwanted network traffic, to high-level information like malware or TTPs analysis, that can benefit security analysts in their investigations. The level of intelligence impacts the feed format, detailed in 3.2.4. There are free and paid solutions, some more reliable than others, in the various existing formats.

Below are some of the widely adopted, free and community-built feeds, that can be imported and integrated in CTI platforms, containing IoCs from all the pyramid levels (detailed in 3.2.3):

- https://www.abuse.ch/
- https://www.anomali.com/resources/limo
- https://www.blocklist.de/en/export.html
- https://www.botvrij.eu/
- https://www.cinsarmy.com/
- http://www.covert.io/threat-intelligence/
- https://iplists.firehol.org/
- https://www.misp-project.org/feeds/
- https://threatfeeds.io/

3.1.8 TIPs - Platforms

ENISA defines TIP as ”emerging technology discipline that supports organisations” [34] threat intelligence programs and helps them to improve their cyber threat intelligence capabilities that enable organisations to easily bootstrap the core processes of collecting, normalising, enriching, correlating, analysing, disseminating and sharing of threat related information”.

According to the SANS CTI Survey 2021 [8], 39.3% of the respondents have single or shared analysts working on CTI and 44.4% have a formal dedicated team focused on CTI, meaning that 83.7% handle CTI and takes profit of it (see Figure 3.8).

Threat intelligence cannot be effective without tools that can help analysts to look wider and find connections between data., TIPs play an important role there: 71.7% of the respondents integrate CTI into their defense and response systems using TIPs (commercial or open source) (see Figure 3.9).

Figure 3.8: ”Does your organization have resources that focus on CTI?”, extracted from [8]

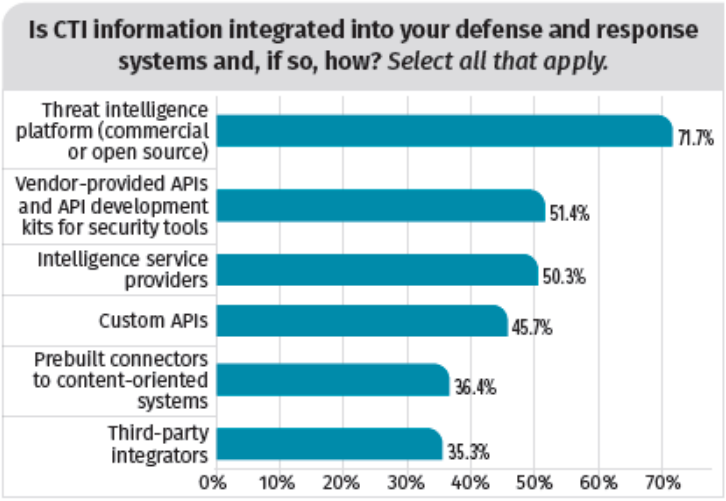


Figure 3.9: ”Is CTI integrated into your DR systems and, if so, how?”, extracted from [8]

The quality of the collected data and intelligence is a major topic but with high levels of dissatisfaction and still with little relevance on TIPs, as discussed in Section 1.1, and it is not normalized and standardized across vendors, platforms and users. The UK NCSC suggests [31] the adoption of the NATO System (also known as Admiralty Grading System - see Appendix A) to evaluate CTI reliability and credibility, by assigning a two-character notation that appraises the reliability of the source, and the information's assessed level of credibility, but that does not take into account other important metrics, as relevance and timeliness.

There are several tools available in the market (Table 3.2), either commercial or open-source. Some vendors also provide free-tier cloud-community solutions that cover a good part of the pyramid of pain, giving access to additional components with IAP or subscription models.

According to ENISA [34], there are some challenges and limitations related to the current state and usage of TIPs:

- the volume of shared CTI is enormous.
- diversity of data models and formats.
- limited workflow, analytics and automation capabilities.
- sharing is focused on less important IoCs.
- focus on collection phase of lifecycle.
- limited enablement in triage and relevancy determination.
- quality trust issues.
- no TTL for shared CTI.

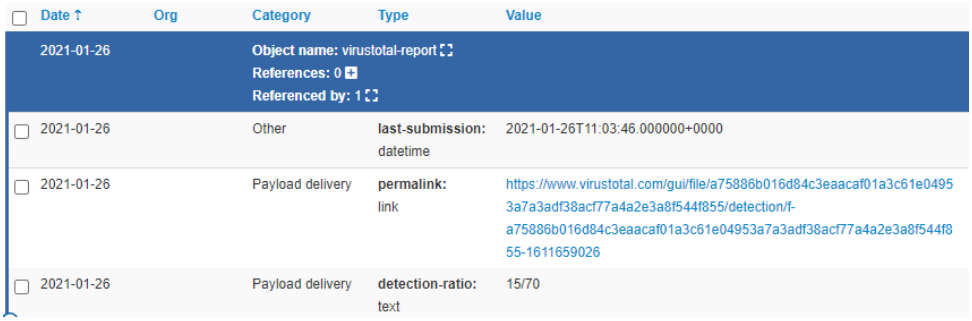
Name	Type	Owner
X-Force Exchange	Community	IBM
Open Threat Exchange (OTX)	Community	AlienVault
OpenCTI	Open-Source	Luatix
SpiderFoot HX	Open-Source	SpiderFoot
threatnote.io	Open-Source	threatnote.io
MineMeld	Open-Source	Palo Alto
Malware Information Sharing Platform (MISP)	Open-Source	CIRCL
scoutTHREAT	Commercial	Lookingglass
ThreatConnect TIP	Commercial	ThreatConnect
Threatstream	Commercial	Anomali
ThreatQ Platform	Commercial	ThreatQuotient
RecordedFuture Platform	Commercial	RecordedFuture
AutoFocus	Commercial	Palo Alto
EclecticIQ TIP	Commercial	EclecticIQ

Table 3.2: Threat Intelligence Platforms

3.2 IoC

3.2.1 Definition

Indicator of Compromise (IoC) is an forensic artifact containing evidences of a certain potential cyberattack that has taken place, providing valuable data obtained from the attacked/target asset. It can contain payloads, network activity and other useful information, so that analysts can decide accordingly (see Figure 3.10). Cyber Security community and top tier companies SOCs collaborate by exchanging IoCs with the purpose of accelerating Cyber Threat detection and response.



Date	Org	Category	Type	Value
2021-01-26		Object name: virustotal-report References: 0 Referenced by: 1		
2021-01-26		Other	last-submission: datetime	2021-01-26T11:03:46.000000+0000
2021-01-26		Payload delivery	permalink: link	https://www.virustotal.com/gui/file/a75886b016d84c3eaacaf01a3c61e04953a7a3adf38acf77a4a2e3a8f544f855/detection/f-a75886b016d84c3eaacaf01a3c61e04953a7a3adf38acf77a4a2e3a8f544f855-1611659026
2021-01-26		Payload delivery	detection-ratio: text	15/70

Figure 3.10: IoC at MISP platform

3.2.2 Lifecycle

IoCs may origin from organizations SOCs, feeds (open or paid – produced by public entities or vendors) or can be created by security researchers/enthusiasts. An IoC object can be extended and enriched during its lifecycle, giving life to other IoCs, i.e. it is not a static and atomic object. The IoC lifecycle is represented in Figure 3.11.

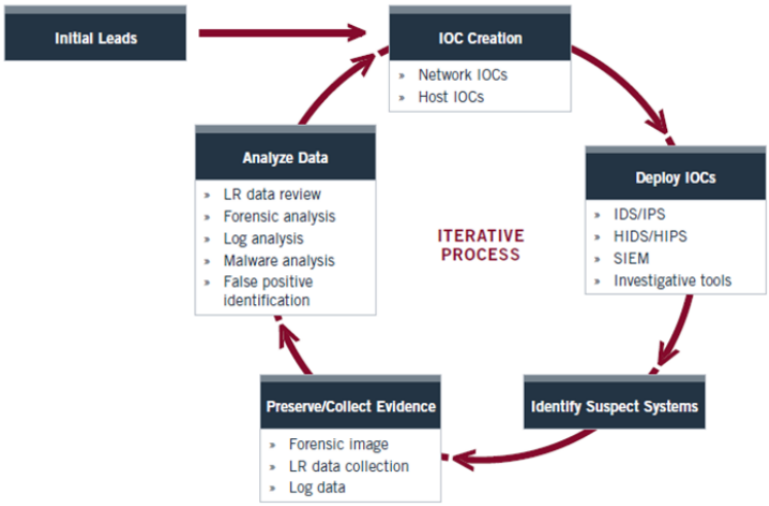


Figure 3.11: IoC lifecycle, extracted from [16]

3.2.3 Types

The ”Pyramid of Pain” [17], represented in Figure 3.2.3, measures the potential usefulness of the IoC and the difficulty of obtaining it, ordered from the least painful and easy to use handle (attack or to defense) to the most difficult ones. The higher, the worse.

From the bottom to the top of the pyramid:

- **Hash values:** MD5, SHA1, SHA256 or other common hash values, used to identify and sign malicious files

Figure 3.12: The Pyramid of Pain - different IoC types, extracted from [17]

- **IP Addresses:** Typically ”Command and control” (C2, C&C) malicious IPs
- **Domain Names:** Domains and sub-domains used by attackers
- **Network/Host Artifacts:** Footprints, traces and artifacts left by attackers. On hosts, abnormal files/directories or registry objects. On the network side, protocol errors or typos.
- **Tools:** Tools used by attackers left fingerprints, either by its behaviour or by some specific characteristic, that can be used by detection mechanisms to identify and response. Some of these tools are identified and listed in the MITRE ATT&CK framework.
- **TTPs:** Attacker operation description, containing all the ”Tactics, Techniques and Procedures”

3.2.4 Formats

Sharing with partners, peers and the community is the key of success. When dealing with huge volumes of complex information, it is important to represent it in a standardized and structured format, ideally expressive, flexible, extensible, automatable, and as human-readable as possible. Besides the common CSV format, some structured and open-source formats have been developed across the community:

- **MISP**: A JSON based format, MISP core was developed by the CIRC Luxembourg to be the standard format for the exchange of events and attributes across MISP instances, as it was designed to support other implementations which reuse the format and ensuring an interoperability with existing MISP support. Its meaning only

depends of the information embedded. An event can be originated in an incident, a security analysis report or a specific threat actor analysis and is composed by:

  - Event attributes: containing high level information, like the unique identifier, threat level, timestamp, date
  - Objects: with both the source (Orgc) and originator (Org) organizations identification
  - Attribute: represented in a category-type-value triplet, it is used to describe the indicators and contextual data of an event, where the category and type give meaning and context to the value.
  - ShadowAttribute: 3rd party attributes that either propose to add new or modify information to an event.
  - Object: contextual bond between a list of attributes within an event, i.e. allows the description of complex structures than can be described by a single attribute, like files.
  - Object References: serves as a logical link between an Object and other referenced Object or Attribute.
  - EventReport: used to complement an event with one or more report.
  - Tag: to classify an event with a freely chosen string.
  - Sighting: describes whether an attribute has been seen under a given set of conditions.
  - Galaxy: clustering of events.

- STIX: originally developed by MITRE, Structured Threat Information eXpression (STIX) is a structured language and serialization format used to describe CTI, so it can be shared, stored and analyzed consistently.

  Currently maintained by the CTI Technical Committee at OASIS and in v2.1, STIX is one of the most adopted formats, as it supports several (currently 18) distinct domain objects (see Figure 3.14), that represent CTI unique concepts, for situational awareness, real-time network defense, and sophisticated threat analysis:

  - Attack Pattern: TTP description, including adversaries attempts to compromise targets.
  - Campaign: description of a set of malicious activities or attacks, that occur over a period of time against a specific set of targets.
  - Course of Action: set of recommendation that might be taken in response to a certain intelligence.

```
{
  "Event": {
    "uuid": "0b988513-9535-42f0-9ebc-5d6aec2e1c79",
    "timestamp": "1607324075",
    "info": "OSINT - Egregor: The New Ransomware Variant To Watch",
    "date": "2020-11-27",
    "published": true,
    "extends_uuid": "",
    "publish_timestamp": "1607324084",
    "analysis": "2",
    "threat_level_id": "1",
    "Orgc": {
      "uuid": "55f6ea5e-2c60-40e5-964f-47a8950d210f",
      "name": "CIRCL"
    },
    "Tag": [
      {
        "colour": "#004646",
        "name": "type:OSINT"
      },
      {
        "colour": "#0071c3",
        "name": "osint:lifetime=\"perpetual\""
      },
      {
        "colour": "#0087e8",
        "name": "osint:certainty=\"50\""
      },
      {
        "colour": "#ffffff",
        "name": "tlp:white"
      },
      {
        "colour": "#0088cc",
        "name": "misp-galaxy:ransomware=\"Egregor\""
      }
    ],
    "Attribute": [
      {
```


Figure 3.13: MISP core format - sample, extracted from [18]

- Grouping: allows to set a shared context with other STIX objects.
- Identity: individuals, organizations, or groups identified as actors (either targets or attackers).
- Indicator: pattern that ca be used to identify and detect suspicious activity.
- Infrastructure: describes any systems, software or resource that constitutes the infrastructure, i.e. the attack surface.
- Intrusion Set: behaviors and resources with common properties, that can link to a single organization.
- Location: represents a geographic location, either a city, region, address, latitude and longitude.
- Malware: characterizes, identifies, and categorizes a specific instance of iden-

tified malware.

- Malware Analysis: includes the metadata and results of the analysis performed on a identified malware.
- Note: allows to insert extra information not contained in other STIX Objects, to provide further context and/or additional analysis.
- Observed Data: provides information about related entities such as files, systems, and networks using SCOs.
- Opinion: assessment produced by a different entity about the correctness of the information included in a certain SDO.
- Report: TI collection about one or more topics, such as threat actor, malware or attack technique.
- Threat Actor: describes the individuals, groups, or organizations believed to be behind the malicious activity.
- Tool: characterizes the software tools used by threat actors to perform attacks.
- Vulnerability: used to link to external definitions of vulnerabilities or to describe 0-day vulnerabilities.

These objects can be linked through relationships or sightings, the two existing connectors in STIX, as represented in the Figure 3.15.

- Relationship: allows to link and describe the relation of two SDOs.
- Sighting: used to link a SDO with who sighted and/or it was sighted, and what was actually sighted (Observed data).

- OpenIOC: originally developed by Mandiant in 2011 and currently maintained by FireEye (that acquired the company and keeps the brand as a product), Open Indicators of Compromise (OpenIOC) is based on XML and it enables the description of characteristics that either identify a threat or attack methodology. It is not as modular and expressive as STIX but designed to be simple, effective and easy to read.
- RSS/TXT/CSV: free text, separated by lines/commas/semicolons/tabs/pipes, usually used for IP/DNS/URL/hash blacklists. This is the simplest and fastest format to maintain and parse, with no more details or information included and it is frequently adopted to block, ignore or drop communication to/from undesired actors.

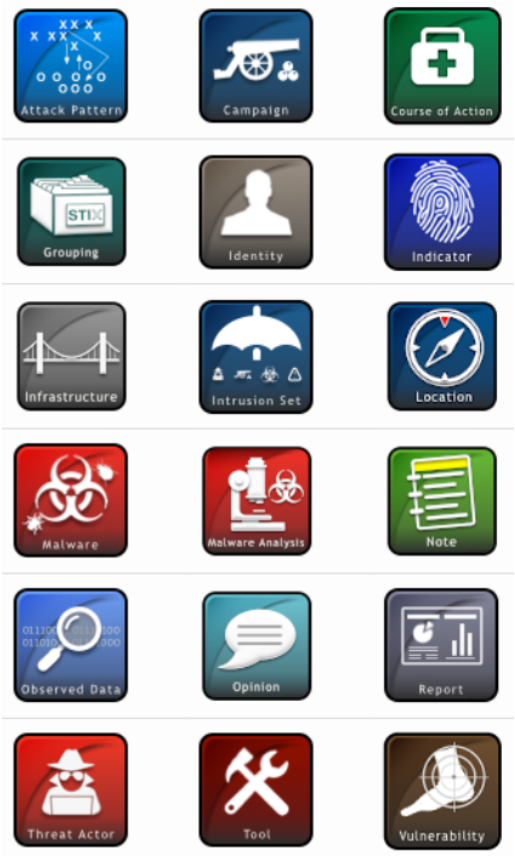


Figure 3.14: STIX - domain objects, extracted from [19]

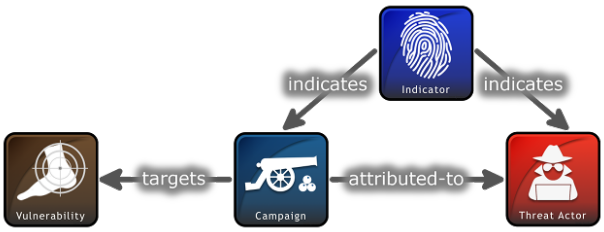


Figure 3.15: STIX - relationship example, extracted from [19]

```
{
    "type":"bundle",
    "id":"bundle--2a25c3c8-5d88-4ae9-862a-cc3396442317",
    "objects":[
        {
            "type":"indicator",
            "spec_version":"2.1",
            "id":"indicator--a932fcc6-e032-476c-826f-cb970a5a1ade",
            "created":"2014-02-20T09:16:08.989Z",
            "modified":"2014-02-20T09:16:08.989Z",
            "name":"File hash for Poison Ivy variant",
            "description":"This file hash indicates that a sample of Poison Ivy is present.",
            "indicator_types":[
                "malicious-activity"
            ],
            "pattern":"[file:hashes.'SHA-256' = 'ef537f25c895bfa782526529a9b63d97aa631564d5d789c2b765448c8635fb6c']",
            "pattern_type":"stix",
            "valid_from":"2014-02-20T09:00:00Z"
        },
        {
            "type":"malware",
            "spec_version":"2.1",
            "id":"malware--fdd60b30-b67c-41e3-b0b9-f01faf20d111",
            "created":"2014-02-20T09:16:08.989Z",
            "modified":"2014-02-20T09:16:08.989Z",
            "name":"Poison Ivy",
            "malware_types":[
                "remote-access-trojan"
            ],
            "is_family":false
        },
        {
            "type":"relationship",
            "spec_version":"2.1",
            "id":"relationship--29dcdf68-1b0c-4e16-94ed-bcc7a9572f69",
            "created":"2020-02-29T18:09:12.808Z",
            "modified":"2020-02-29T18:09:12.808Z",
            "relationship_type":"indicates",
            "source_ref":"indicator--a932fcc6-e032-476c-826f-cb970a5a1ade",
            "target_ref":"malware--fdd60b30-b67c-41e3-b0b9-f01faf20d111"
        }
    ]
}
```


Figure 3.16: STIX - Malware IoC sample, extracted from [20]


```
<?xml version="1.0" encoding="UTF-8"?>
<ioc xmlns="http://schemas.mandiant.com/2010/ioc" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="0294496a-b037-55b9-a3fe-46a344d7f524" last-
modified="2016-11-08T16:43:21.4083202Z">
   <short_description>ioc for 495DB359D61411F0688211C8DD473CB7</short_description>
   <description>ZONE:Yellow</description>
   <authored_by>KasperskyThreatLookup</authored_by>
   <authored_date>2016-11-08T16:43:21.4083202Z</authored_date>
   <links />
   <definition>
      <Indicator operator="OR" id="6f7a75ee-f423-5cf1-ad42-140ae6aa2301">
         <IndicatorItem condition="is" id="59b35d49-14d6-f011-6882-11c8dd473cb7">
            <Context document="FileItem" search="FileItem/Md5sum" type="mir" />
            <Content type="md5">495DB359D61411F0688211C8DD473CB7</Content>
         </IndicatorItem>
      </Indicator>
   </definition>
</ioc>
```


Figure 3.17: OpenIOC - example, extracted from [21]

Chapter 4

Metrics and dimensions for assessing CTI

This chapter introduces and describes our proposal of metrics for assessing CTI, namely ”Big Data” quality and IoC quality, that corresponds to Twitter account and Tweet assessment, respectively, that we will use to highlight and grant that certain indicator must have more attention and priority than others. We provide an overview of the assessment integration in the DiSIEM OTD [4] platform and the way to view and manage it.

4.1 (Big) Data quality - Twitter account

A large number of data quality dimensions were proposed in the reviewed literature, in Chapter 2, that we consider relevant to score Twitter accounts and its messages, known as ”Tweets”. Ensuring data quality is quite important, and it takes even more relevance when it comes to protecting data, networks, assets and people. This is addressed in ISO/IEC 25012:2008 [5], where it is defined a general data quality model for data retained in a structured format within a computer system. Fifteen quality characteristics are defined for data, categorized according to two points of view:

- **Inherent data quality:** refers to quality characteristics that measure the data itself;
- **System dependent data quality:** quality characteristics that measure the capabilities of the computer system retaining the data.

Some characteristics are relevant in both views, as presented in Table 4.1:

Characteristic	Inherent	System dependent
Accuracy	X	
Completeness	X	
Consistency	X	
Credibility	X	
Currentness	X	
Accessibility	X	X
Compliance	X	X
Confidentiality	X	X
Efficiency	X	X
Precision	X	X
Traceability	X	X
Understandability	X	X
Availability		X
Portability		X
Recoverability		X

Table 4.1: Data quality characteristics defined in ISO/IEC 25012:2008

The mutually exclusive ”System dependent” characteristics (Availability, Portability, Recoverability) are not relevant for our work, even if they are very important in other contexts.

As it is a common and relevant research topic, several studies [35][36][37] analyzed it for Big Data contexts, where the usual data quality dimensions are not so applicable. That said, and also taking the ISO/IEC 25012:2008 as inspiration, we propose a set of dimensions that we consider to be the core ones that must be taken into account to assess Twitter accounts, presented in Table 4.2.

D	Dimension	Description	Source
1	Maintenance	How often Tweets are consumed	[38]
2	Accuracy	False-positive relation	[38]
3	Verifiability	How often provides source	[38]
4	Intelligence	How often offers external objects	[38]
5	Timeliness	How often creates clusters objects	[38]
6	Completeness	How much represents	[38]
7	Profile	Credibility characteristics	[28]

Table 4.2: Twitter source account - core dimensions

It is important to keep the following metrics up-to-date during the assessment, as they are the basis and essential for the computation of dimensions:

- #SystemTweets (i.e.: the number of tweets gathered and processed by the system, by all accounts) = $\sum Tweets(system)$
- #AccountTweets (i.e.: the number of tweets gathered and processed that were "tweeted" by a certain account) = $\sum Tweets(\alpha)$
- #AccountRelevantTweets (i.e.: the number of relevant tweets gathered and processed that were "tweeted" by a certain account) = $\sum RelevantTweets(\alpha)$
- #AccountSourcedTweets (i.e.: the number of tweets gathered and processed that were "tweeted" by a certain account and that contains a reference to the source) = $\sum SourcedTweets(\alpha)$
- #AcountReferencedTweets (i.e.: the number of tweets gathered and processed that were "tweeted" by a certain account and that contains a reference to an external object) = $\sum ReferencedTweets(\alpha)$
- #AccountCreatedClusters (i.e.: the number of clusters (group of tweets that relate to the same IoC) created by a certain account, which happens when that account is the fastest to tweet about a certain topic) = $\sum CreatedClusters(\alpha)$
- #AccountParticipatedClusters (i.e.: the number of clusters where a certain account participates, representing the diversity of topics tweeted by that account)

  = $\sum ParticipatedClusters(\alpha)$

All the dimensions are measured between 0 and 1. The account α score is the weighted average of all the dimensions [D1..D7]. The weights can be customized but by default we will consider them equally distributed. Below we describe in detail these dimensions and the calculation formula for each Twitter account α score:

- D1 (Maintenance): How often messages are added, i.e. the relative contribution of the account itself for the total number of tweets. Each tweet posted by the account α scores for that account. It is defined as:

$$D1(\alpha) = \frac{\sum Tweets(\alpha)}{\sum Tweets(system)} \tag{4.1}$$

- D2 (Accuracy): How often tweets of a Twitter account are valid ones, i.e., the relative number of relevant tweets. This allow us to measure trustworthiness consistency and it is defined as:

$$D2(\alpha) = \frac{\sum RelevantTweets(\alpha)}{\sum Tweets(\alpha)} \tag{4.2}$$

- D3 (Verifiability): How often a Twitter account verifies the information provided by linking its source in the tweet. If a tweet contains a source/link, we consider it as verifiable and its account α scores. This is calculated as:

$$D3(\alpha) = \frac{\sum SourcedTweets(\alpha)}{\sum Tweets(\alpha)} \tag{4.3}$$

- D4 (Intelligence): How often a Twitter account offers other objects in their messages by linking it (blog articles, papers, conference presentations, or similar). If a tweet contains a second source/link, we consider it as intelligence and the account α scores. This is calculated as:

$$D4(\alpha) = \frac{\sum ReferencedTweets(\alpha)}{\sum Tweets(\alpha)} \tag{4.4}$$

- D5 (Timeliness): Quickness ranking. Each cluster created must score points to the account score. If the tweet is the earliest, i.e creates the cluster, we consider it timely and the account α scores 1 point. This is calculated as:

$$D5(\alpha) = \frac{\sum CreatedClusters(\alpha)}{\sum ParticipatedClusters(\alpha)} \tag{4.5}$$

- D6 (Completeness): How much of the entire system a single Twitter account represents. Each participated cluster scores for the account $alpha$. This is calculated as:

$$D6(\alpha) = \frac{\sum ParticipatedClusters(\alpha)}{\sum Clusters(system)} \tag{4.6}$$

- D7 (Profile): A Twitter profile has many characteristics, like registration date, location, number of followers, verified, URL, description. To simplify the evaluation of the profile we suggest to calculate this metric as:

$$\begin{aligned} D7(\alpha) = {} & \mathbb{1}_{Accounts\ more\ than\ 6\ months\ old}(\alpha) * 1/3 + \\ & \mathbb{1}_{Accounts\ verified}(\alpha) * 1/3 + \\ & \mathbb{1}_{Accounts\ with\ more\ than\ 1000\ followers}(\alpha) * 1/3, \end{aligned} \tag{4.7}$$

  where $\mathbb{1}_A(\alpha)$ is the characteristic function of A, that is, $\mathbb{1}_A(\alpha) = 1$, if $\alpha \in A$, and $\mathbb{1}_A(\alpha) = 0$, otherwise.

And finally, the classification of the Twitter account itself:

- Account score: Weighted average of all the previous described dimensions [D1..D7].

$$Score(\alpha) = \sum_{d=1}^{7} w_d x_d \tag{4.8}$$

where w_d is the weight of dimension d, with $w_d >= 0$ and $\sum_{d=1}^{7} w_d = 1$

4.2 IoC quality - Tweets

We propose that each single Tweet message and IoC also have his own score, where its account must have relevance, and contributing for the score, as stated in the table below:

D	Dimension	Description	Source
8	Validity	Likelihood of containing valid and valuable information	[24]
9	Relevance	IT assets/infrastructure cross-check	[24]
10	Timeliness	Cluster fitness	[24]
11	Source account	Table 1. metrics influence	[28]
12	Tweet features	Likes, retweets, comments	[28]

Table 4.3: Single Tweet message - core dimensions

All the dimensions are measured between 0 and 1. The Tweet τ score is the weighted average of all the dimensions. The weights can be customized but by default we will consider it equally distributed. Below we describe in detail these dimensions and the calculation formula for each one:

- D8 (Validity): Using the binary classifier from MCD [7] (see Chapter 5.1), that represents the likelihood of each tweet containing valid and valuable information about an asset of interest, or not, we set the validity of the tweet, simply as:

$$D8(\tau) = BinaryClassification(\tau) \tag{4.9}$$

- D9 (Relevance): Using the NER (see Table 5.1) thread from both SYNAPSE [9] and MCD [7], we are able to measure how much of the monitored infrastructure is affected (each assets can have custom weight), i.e. if it is relevant for a certain infrastructure or a predefined set of assets.

$$D9(\tau) = \sum_{i}^{n} w_i x_i \quad \text{, where} \quad w_i >= 0 \quad , \sum_{i}^{n} w_i = 1 \tag{4.10}$$

and w_i is the weight/importance of a certain asset i.

- D10 (Timeliness): The inclusion of the tweet inside a cluster will affect its score and it will be worse according to the "inclusion" timestamp, i.e the further away, the lower the score. On the other hand, if the tweet is the root of the cluster, it will have the best possible score for D10, i.e D10(τ) = 1. That said, we propose to define D10 as time conditional:

$$D10(\tau) \begin{cases} 1 & \text{, if } \tau \text{ is the root of the cluster} \\ \frac{100 - \frac{|\Delta(Timestamp(\tau, root))|}{2}}{100} & \text{, otherwise} \end{cases} \tag{4.11}$$

- D11 (Source account): The global score of the tweet source account must have influence on the tweet message score. The goal here is include the trustworthiness of the source account, based mostly on whether it has provided in the past.

$$D11(\tau) = AccountScore(\tau) \quad \text{, as defined in (3.8)} \tag{4.12}$$

- D12 (Tweet features): A tweet has some public metrics, like number of likes, number of retweets/quotes, number of replies/comments. This is only relevant after some time (when the tweet is submitted, the count of likes, retweets/quotes or replies/comments is 0) and if recalculated in time, as the goal here is to have timeliness tweets:

$$\begin{aligned} D12(\tau) = {} & \mathbb{1}_{More\ than\ 10\ likes}(\tau) * 1/ + \\ & \mathbb{1}_{More\ than\ 5\ retweets/quotes}(\tau) * 1/3 + \\ & \mathbb{1}_{More\ than\ 2\ replies/comments}(\tau) * 1/3, \end{aligned} \tag{4.13}$$

  where $\mathbb{1}_A(\tau)$ is the characteristic function of A, that is, $\mathbb{1}_A(\tau) = 1$, if $\alpha \in A$, and $\mathbb{1}_A(\tau) = 0$, otherwise.

And finally both the Tweet and IoC scores:

- Tweet score: Weighted average of all the previous described dimensions [D8..D12].

$$TweetScore(\tau) = \sum_{d=8}^{12} w_d D_d \quad \text{, where} \quad w_d >= 0 \quad \text{and} \sum_{d=8}^{12} w_d = 1 \tag{4.14}$$

- IoC (cluster) score: Maximum "Tweet" score within the cluster.

$$IoCScore(\iota) = \max(TweetScore(\tau_1), ..., TweetScore(\tau_n)) \tag{4.15}$$

All the mentioned metrics are normalized and measured through a value between [0,1].

The next chapter will focus on refining and analysing these dimensions against data extracted in [9], assessing it and calculating a score for each Tweet, considered Twitter account and cluster. The target is to supply SIEM platform and its users with automatic assessment, easing and accelerating the take of decisions by cybersecurity analysts.

Chapter 5

Dataset analysis and data enrichment

In this chapter we describe in detail the pipeline architecture and how the dataset goes from raw data to enriched data, with all the assessments made. Having defined our architecture, we report the experimental evaluations and how the score can vary according to the user needs and the metrics weights.

5.1 Assumptions

The following assumptions must be taken into account:

- Assumption 1: The dataset here analysed (**DS1**) contains data collected on *DiSIEM OTD* [4] between 15/July/2021 and 14/Sep/2021, containing 1081 clusters, 1568 tweets, submitted by 54 different accounts, targeting a cluster of keywords previously selected that represent the real infrastructure of a well known Utilities company. The list of keywords and assets is described in Table 5.8. At this stage, the only assessment provided in the dataset comes from *Within-cluster Threat Similarity (WTS)*, a cohesion measure that relates the number of words shared by all the cluster's tweets and the number of words of the smallest tweets, i.e. WTS is 0 if no words are shared by the cluster's tweets, and 1 when all tweets share the words of the smallest tweet in the cluster.
- Assumption 2: The dataset **DS1** was transformed using the MCD classifier *"multitask-cyberthreat-detection"* [7], so that we were able to enrich the dataset with:
  - Clean text: completely cleaned tweet, without special characters, only words.
  - Entities: summary of found entities like vendors, assets, versions, threats, vulnerability IDs
  - Binary Pred. Confidence: by using a Multi-Task Learning approach, *Multitask Cyberthreat Detection (MCD)* outputs a binary classifier that represents the relevancy of each tweet/sentence, i.e. the likelihood of containing valuable information about an asset of interest, or not.

– Binary Pred.: Binary Pred. Confidence rounded according to a threshold (<0.5 » 0, ≥0.5 » 1).

– Tags: entity labeling, using *Named-entity recognition (NER)* (see Table 5.1)

Label	Description
O	Other - ”Does not contain useful information.”
ORG	Organization - ”Company or organization.”
PRO	Product - ”Product or asset.”
VER	Version - ”A version number, possibly from the identified asset or product.”
VUL	Vulnerability - ”May be referencing the existence of a threat or a vulnerability.”
ID	Identifier - ”An identifier, either from a public vulnerability repository (e.g., NVD) or from an update or patch.”

Table 5.1: Entity labeling - tags

- Assumption 3: The dataset was parsed, loaded and analysed in a Java application made for that purpose. After this work, all the dimensions/metrics can be integrated in a improved version of the DiSIEM OTD [4] platform in order to improve the user experience. Figure 5.1 represents the computation timeline at the different stages of SYNAPSE [9].

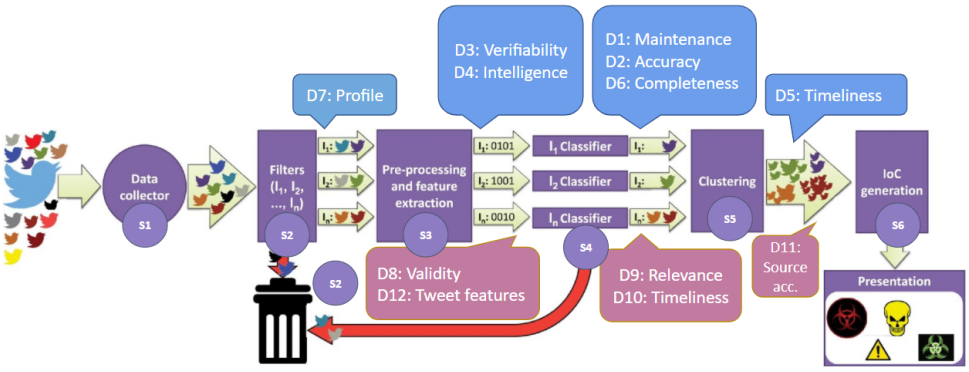


Figure 5.1: Integration of metrics with the SYNAPSE pipeline

- Assumption 4: Following SYNAPSE computation timeline order, we assume that:

  – **after S2 (filtering part)**, it is likely that the tweet will be consumed:
  For **D7**, if the profile is already known by the system, it is required to update it periodically; otherwise gather profile characteristics.

  – **after S3 (pre-processing and feature extraction)**, the tweet is clean, we are able to update both **D3 and D4**.

  - **after S4 (classifier) execution**, we are sure that the tweet will be consumed and we have all the required data to update **D1, D2, D6, D9 and D10**.
  - **after S5 (clustering) execution**, the IoC is ready to be created/updated and both **D5 and D11** can be (re)calculated.

- Assumption 5: The dataset DS1 does not contain details about the Twitter accounts, i.e profile characteristics, neither public metrics about the gathered Tweets. This is queried from Twitter API v2, to which we apply for access, since access is provided for Academic Research projects, using the following queries:

  - **Q1: GET /2/users/by/username/:username** - Account API

    This request allows us to obtain more information about the Twitter account, like the location and the public metrics, by passing the path and query parameters defined in Tables 5.2 and 5.3.

Name	Type	Description
username	string	The Twitter username (handle) of the user

Table 5.2: Path parameters - Q1

Name	Type	Description
user.fields	enum	Where we specify the desired fields, in this case the ones mentioned in Table 5.4

Table 5.3: Query parameters - Q1

Name	Type	Description
created_at	date	Account creation timestamp
description	string	The text of this user's profile description, also known as bio(graphy)
location	string	The location specified in the user's profile
name	string	The friendly name of this user, as shown on their profile
public_metrics.followers_count	integer	Number of users who follow this user.
public_metrics.following_count	integer	Number of users this user is following.
public_metrics.tweet_count	integer	Number of Tweets (including Retweets) posted by this user.
verified	boolean	Indicate if this user is a verified Twitter user

Table 5.4: Response fields - Q1

– **Q2: GET /2/tweets/:id** - Tweet API

This request allows us to complement Tweets information with data that is not present in the dataset, like the public metrics (see Table 5.7), by passing the path and query parameters defined in Tables 5.5 and 5.6.

Name	Type	Description
id	string	Tweet unique identifier to request

Table 5.5: Path parameters - Q2

Name	Type	Description
tweet.fields	enum	Where we specify the desired fields, in this case the public metrics mentioned in Table 5.7

Table 5.6: Query parameters - Q2

Name	Type	Description
public_metrics.retweet_count	integer	Number of times this Tweet has been Retweeted.
public_metrics.reply_count	integer	Number of Replies of this Tweet.
public_metrics.like_count	integer	Number of Likes of this Tweet.
public_metrics.quote_count	integer	Number of times this Tweet has been Retweeted with a comment (also known as Quote).

Table 5.7: Response fields - Q2

These two queries (Q1 and Q2) contribute to both D7 and D12, having the ”Account API” gathering profile data and the ”Tweet API” public metrics for tweets. The Twitter API v2 offers other interesting metrics, namely the *”non public metrics”* and *”organic metrics”*, that contain information like the number of visualizations of a certain tweet, that are not used in the computation of the defined metrics, therefore not taken into account in this work.

It is also worth mentioning that as an *”Academic Research product”*, we are rate-limited, which means that both queries are limited to 900 requests per 15-minute window. This constraint can slow down the update of data through Twitter and therefore the update of the metrics/dimensions, i.e., it can trigger a bottleneck. In the limit, if we talk about huge amounts of data, it can interfere with the timeliness dimension. Nonetheless, this could be mitigated with the use of Twitter Premium or Enterprise grade subscriptions.

Asset	Keywords
.NET framework	.net
Adobe Acrobat Reader	acrobat reader
Adobe Shockwave	shockwave;flash
Apple Quicktime	quicktime
Chrome browser	google chrome;chrome
Cute PDF writer	cutepdf;cute pdf
IBM AIX OS	aix
IE browser	internet explorer;iexplorer
Java	java;jre
Linux	linux
McAfee AV	mcafee
Microsoft Exchange	microsoft exchange;ms exchange;exchange
MS Office Suite	microsoft office;ms office;office;word;excel;powerpoint
MS Silverlight	silverlight
MS Visio	visio
Prisma Cloud	prisma cloud
Probely	probely
Red Hat Enterprise Linux OS	rhel;red hat;redhat
SAP software	SAP
SCADA systems	scada
Solarwinds Orion	solarwinds;solarwinds orion;sunburst;solorigate
Windows 10 OS	windows10;windows 10;win10
Windows 2000 OS	windows 2000;win2000;win 2000
Windows 2003 OS	windows 2003;win2003;win 2003
Windows 2008 OS	windows 2008;win2008;win 2008
Windows 2012 OS	windows 2012;win2012;win 2012
Windows 2016 OS	windows 2016;win2016;win 2016
Windows 7 OS	windows7;windows 7;win7;win 7
Windows 8.1 OS	windows 8.1;win8.1;win 8.1
Windows NT OS	windows nt;winnt;win nt
Windows XP OS	windows xp;winxp;win xp

Table 5.8: Assets and keywords adopted

5.2 Architecture

The architecture represented in Figure 5.2 was adopted due to the usage of both phases 1 and 2, where two existing and mature LASIGE projects (DiSIEM OTD [4] and MCD [7]) deal with the Twitter data collection and analysis. The product of this project, represented in phase 3, outputs CSV files that can be used to create a shared dashboard with the IoC analytics, or to integrate into a customized SIEM input that is customized to accept it as input. It is also prepared to output both MISP and STIX formats, with all the data, plus the new fields introduced by phase 2 and 3.

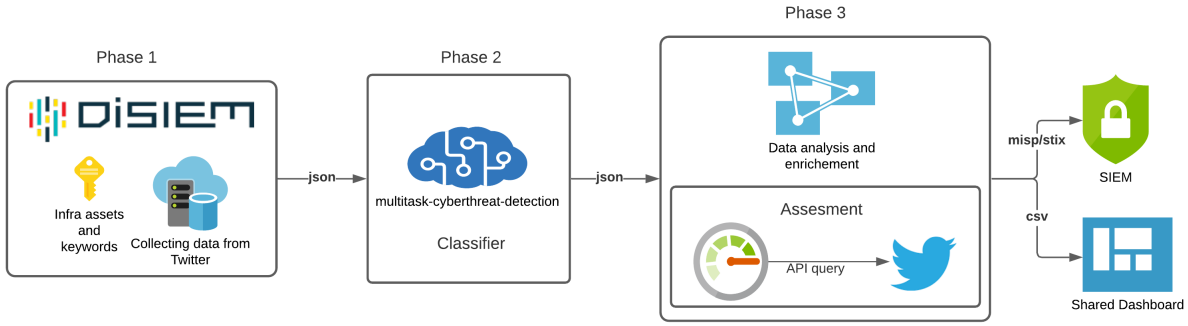


Figure 5.2: Architecture

Phase	Tool	Output format	Output description
Phase 1	DiSIEM OTD	json	Raw data from DiSIEM OTD dump, containing all Tweets collected. Extracted using elasticdump.
Phase 2	MCD	json	DiSIEM OTD dump transformed by MCD.
Phase 3	Java app.	csv	To create shared dashboard with data assessment.
Phase 3	Java app.	MISP/STIX	This output can be developed so that IoCs can be integrated in any SIEM platform that support either MISP or STIX formats.

Table 5.9: Different phases outputs

Taking a simple cluster from **DS1** as an example, we can observe in Figures 5.3 and 5.4, the differences between the outputs of both phases 1 and 2. The output from phase 2 is the output from phase 1, with the bonus provided by the MCD classifier *"multitask-cyberthreat-detection"* [7], that introduces several fields for each Tweet of the cluster, as mentioned in Section 5.1.

```
{
    "_index":"clusters_francisco",
    "_type":"clusters",
    "_id":"1425589191669424130",
    "_score":1,
    "_source":{
        "@timestamp":"2021/08/11 22:45:24",
        "size":2,
        "lastUpdated":"2021/08/11 22:45:24",
        "exemplar":"CVE-2021-1112 NVIDIA Linux kernel distributions contain
      a vulnerability in nvmap, where a null pointer dereference may lead
      to complete denial of service. https://t.co/piJlcoGbgM",
        "threatId":[
            "CVE-2021-1112"
        ],
        "hidden":false,
        "exemplarLinks":[
            " https://t.co/piJlcoGbgM"
        ],
        "wts":0.75,
        "version":2,
        "threatIds":[
            "CVE-2021-1112",
            "CVE-2021-1114"
        ],
        "keywords":[
            "linux"
        ],
        "account":"CVEnew",
        "relevant":true,
        "tweets":[
            {
                "hidden":false,
                "text":"CVE-2021-1114 NVIDIA Linux kernel distributions contain
            a vulnerability in the kernel crypto node, where use after free
            may lead to complete denial of service. https://t.co/3PXTECuHq7",
                "links":[
                    " https://t.co/3PXTECuHq7"
                ],
                "id":"1425589195406462985",
                "date":"2021/08/11 22:45:25",
                "account":"CVEnew",
                "relevant":true
            }
        ],
        "threats":[
            "denial of service",
            "vulnerability"
        ]
    }
}
```


Figure 5.3: Phase 1 - Output example

```
{
    "_index":"clusters_francisco",
    "_type":"clusters",
    "_id":"1425589191669424130",
    "_score":1,
    "_source":{
        "@timestamp":"2021/08/11 22:45:24",
        "size":2,
        "lastUpdated":"2021/08/11 22:45:24",
        "exemplar":"CVE-2021-1112 NVIDIA Linux kernel distributions contain
      a vulnerability in nvmap, where a null pointer dereference may lead
      to complete denial of service. https://t.co/piJlcoGbgM",
        "threatId":[
            "CVE-2021-1112"
        ],
        "hidden":false,
        "exemplarLinks":[
            " https://t.co/piJlcoGbgM"
        ],
        "wts":0.75,
        "version":2,
        "threatIds":[
            "CVE-2021-1112",
            "CVE-2021-1114"
        ],
        "keywords":[
            "linux"
        ],
        "account":"CVEnew",
        "relevant":true,
        "tweets":[
            {
                "hidden":false,
                "text":"CVE-2021-1114 NVIDIA Linux kernel distributions contain a
            vulnerability in the kernel crypto node, where use after free may
            lead to complete denial of service. https://t.co/3PXTECuHq7",
                "links":[
                    " https://t.co/3PXTECuHq7"
                ],
                "id":"1425589195406462985",
                "date":"2021/08/11 22:45:25",
                "account":"CVEnew",
                "relevant":true,
                "clean_text":"cve-2021-1114 nvidia linux kernel distributions contain
            a vulnerability in the kernel crypto node where use after free may lead
            to complete denial of service",
                "tags":"B-ID O B-PRO I-PRO O O O O O O B-PRO O O O O O O O O O B-VUL I-VUL I-VUL",
                "entities":{
                    "Company":"",
                    "Asset":"linux kernel",
                    "Version":"",
                    "Threat":"denial of service",
                    "IDs":"cve-2021-1114"
                },
                "binary_pred_confidence":"0.98213863",
                "binary_pred":1
            }
        ],
        "threats":[
            "denial of service",
            "vulnerability"
        ]
    }
}
```


Figure 5.4: Phase 2 - Output example

At the output level, phase 3 produces three distinct CSV files:

1. **Accounts:** file containing all the known accounts, its metrics [D1-D7] and its scores. It can be useful to tune the list of accounts that the DiSIEM OTD platform is following on Twitter, or to just create a dashboard with the best accounts and look at its tweets.

    - Column 1: Account
    - Column 2: Maintenance Calc (D1)
    - Column 3: Maintenance Count (M)
    - Column 4: Accuracy Calc (D2)
    - Column 5: Accuracy Count (A)
    - Column 6: Verifiability Calc (D3)
    - Column 7: Verifiability Count (V)
    - Column 8: Intelligence Calc (D4)
    - Column 9: Intelligence Count (I)
    - Column 10: Timeliness Calc (D5)
    - Column 11: Timeliness Count (T)
    - Column 12: Completeness Calc (D6)
    - Column 13: Completeness Count (C)
    - Column 14: Profile (D7)
    - Column 15: AccountScore

2. **Tweets:** file containing all the known tweets, its metrics [D8-D12] and its scores.

    - Column 1: ID (Tweet unique identifier)
    - Column 2: Tweet
    - Column 3: Account
    - Column 4: Validity (D8)
    - Column 5: Relevance (D9)
    - Column 6: Timeliness (D10)
    - Column 7: Source account (D11)
    - Column 8: Tweet features (D12)
    - Column 9: TweetScore
    - Column 10: BMCD (Binary confidence from MCD)

3. **IoCs:** file containing all the clusters and its scores.

    - Column 1: Size (cluster size)
    - Column 2: Threat list
    - Column 3: Threat IDs list (CVE identifiers)
    - Column 4: Keywords (Asset list)
    - Column 5: Links list
    - Column 6: Last update timestamp
    - Column 7: Tweet exemplar
    - Column 8: Score (IoC score)
    - Column 9: WTS (from DiSIEM OTD)
    - Column 10: BMCD (Binary confidence from MCD)

The application is also prepared to output in MISP/STIX format, so that these enriched IoCs can be integrated in any compatible SIEM platform, and take profit of the entire chain of data assessment and its metrics.

5.3 Graphic interface proposal

In this section we analyse the required improvements for the DiSIEM OTD web interface that would allow the adoption and integration of the proposed metrics, directly on its dashboard:

- IoC: selecting an IoC/cluster would expand its composition and present its metrics, along with a descriptive annotation.
- Account: placing the mouse pointer over or selecting an account name would present a popup with its metrics.
- Tweet: placing the mouse pointer over or selecting a tweet would present a popup with its metrics.
- Weights customization panel: accessible from a new dedicated icon on the right of the ”Assets” tab, this popup would allow the customization of the metrics weights for both the Account and IoC score computation.

Figure 5.5 shows the web interface of the current version, where no data quality assessment is provided to the user.

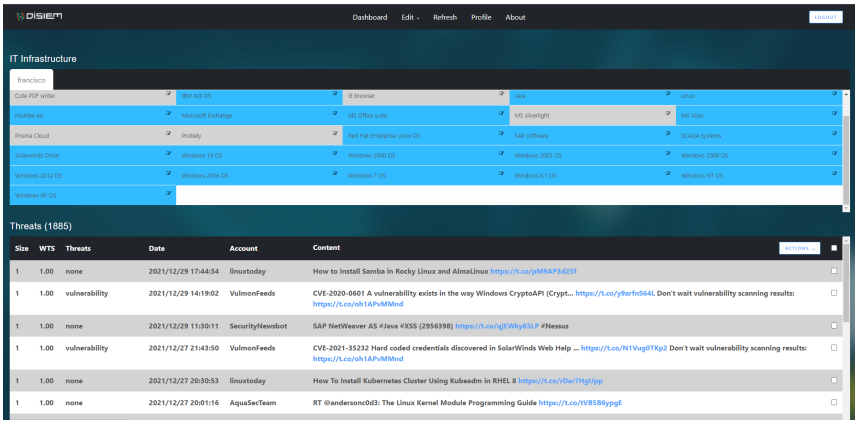


Figure 5.5: DiSIEM OTD web interface - current version

The center of Figure 5.6 contains the actual web interface of the DiSIEM OTD platform, with a selected IoC/cluster, for which we propose to introduce both the IoC metrics and IoC score, located in the area just below the IoC/cluster expanded area, as Figure 5.7 details.

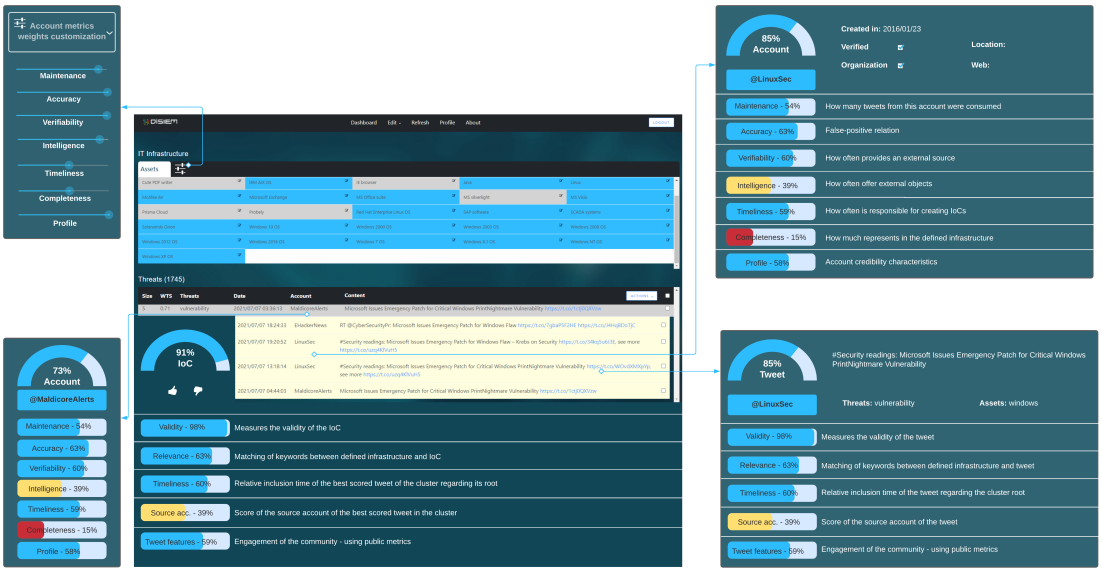


Figure 5.6: DiSIEM OTD web interface proposal

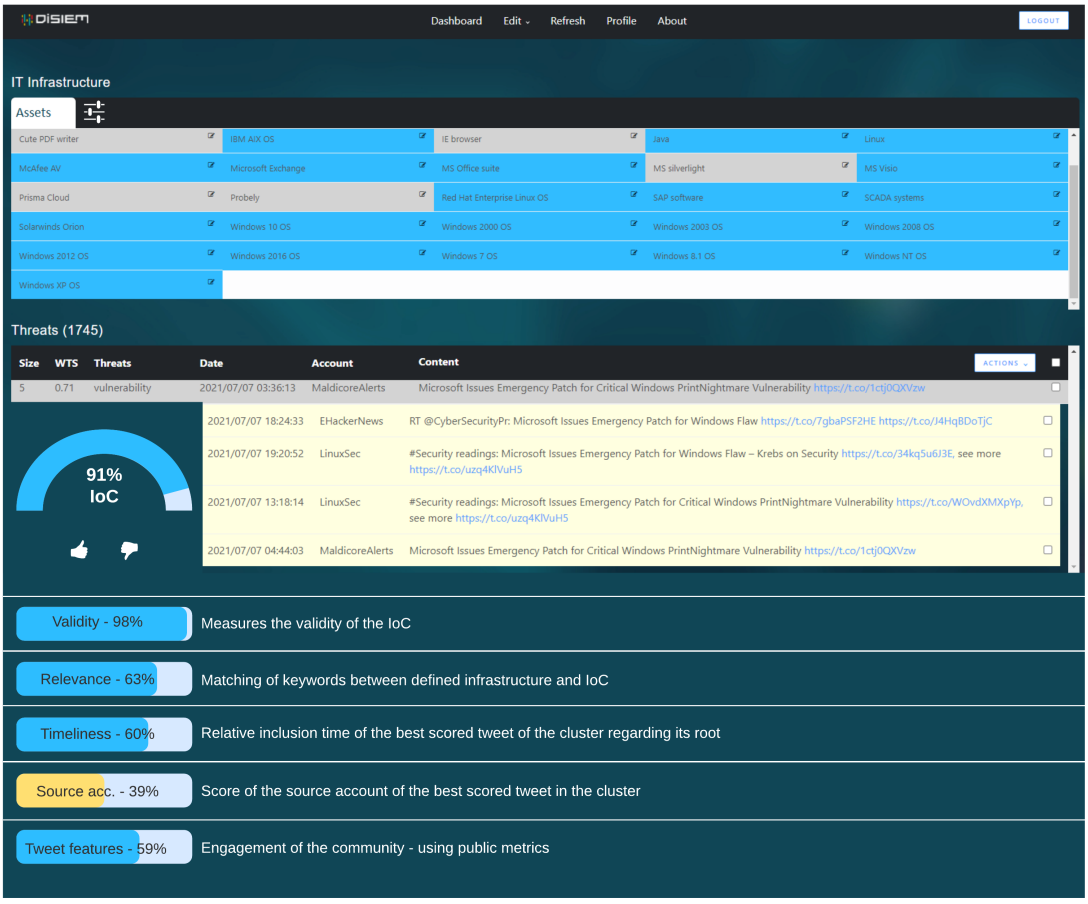


Figure 5.7: DiSIEM OTD web interface - main dashboard with IoC metrics

On the bottom left and top right of Figure 5.6 are represented both the tiny (Figure 5.8) and expanded (Figure 5.9) versions of the account popup panel. The expanded popup contains a descriptive annotation about the metrics and also account details like creation date, location, if verified, if belongs to an organization and the website.

On the bottom right of Figure 5.6 is the tweet metrics panel, detailed below in the Figure 5.10, containing the tweet score, its metrics and a descriptive annotation, the tweet itself and the identified threats and impacted assets.

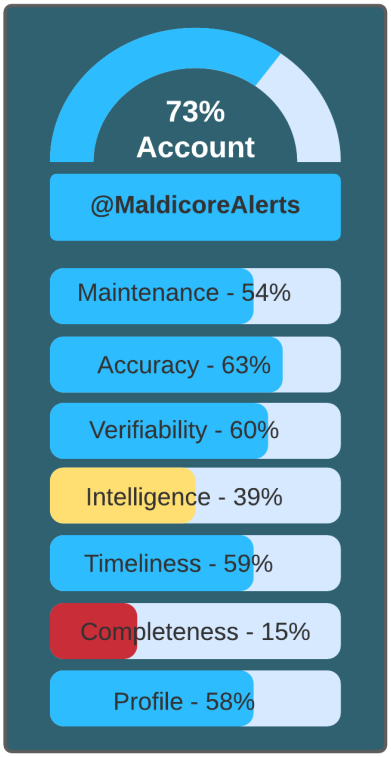


Figure 5.8: DiSIEM OTD web interface - account metrics example, tiny panel

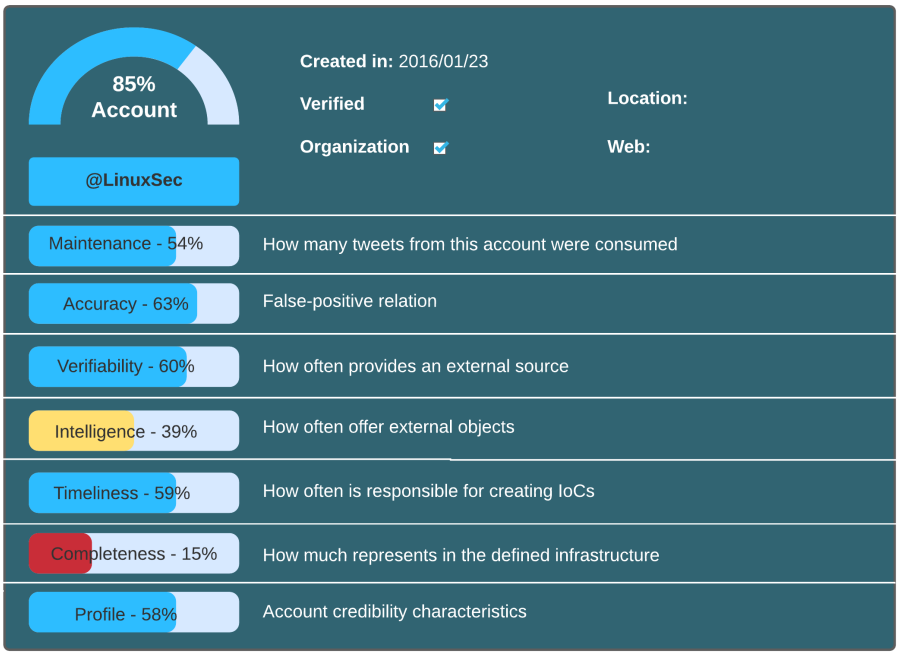


Figure 5.9: DiSIEM OTD web interface - account metrics example, expanded panel

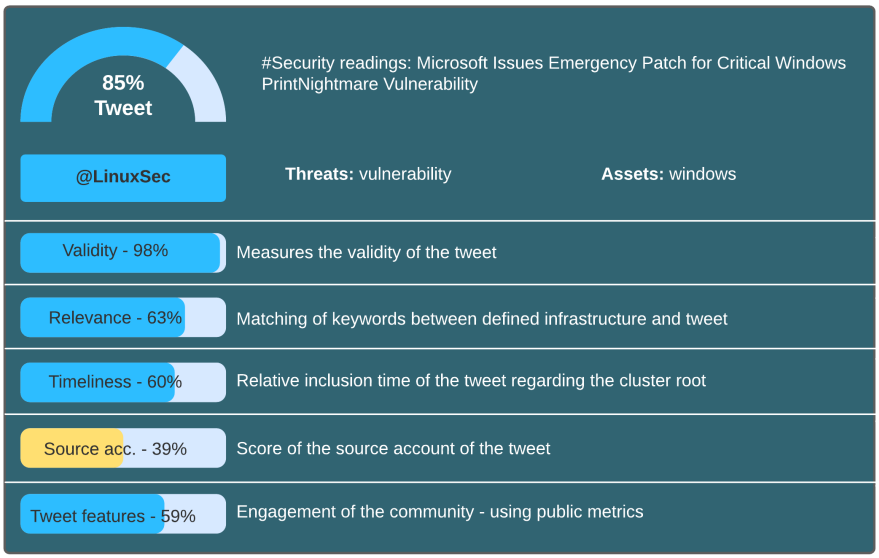


Figure 5.10: DiSIEM OTD web interface - tweet metrics example

The weights customization panel, located in the top left of Figure 5.6 and detailed below in Figure 5.11, is accessible from a new dedicated icon on the right of the ”Assets” tab and allows the users to change the weight of each of the account and twitter metrics, thus influencing the IoC, tweet and account scores, not only because modifying the weight has impact in the way the scores are calculated but also because the account score is one of the tweet score metrics, which in turn also influences the IoC score.

Figure 5.11: DiSIEM OTD web interface - weights customization example

Chapter 6

Experimental Evaluation

From the initial 1081 clusters and 1568 tweets, we filtered out clusters with only 1 tweet and end up with 279 clusters and 487 tweets, submitted by 54 distinct accounts. This cleaning was done to reduce the amount of data to treat and handle, as we were limited by the free Twitter API rates. Even though, we queried **Q2** API 487 times (see Table 5.5) and **Q1** API 54 times (see Table 5.2) in about 8 minutes, which gives approximately 68 queries per minute, 127 queries per 15 minutes, out of the 900 allowed per the rate limit.

If more tweets were consumed and integrated in the dataset, we would still have a comfortable buffer until reaching the rate limit, even having some queries that need to be periodically executed, to maintain the public metrics updated (using **Q2** API). The savings brought about by the fact that we already have all the account data from **Q1** API would also help to not be restricted by the rate limit. This restrictions and concerns only happen because we are treating the full dataset and not working integrated and in parallel with DiSIEM OTD, for the sake of this proof of concept. In a production scenario, the 3 phases from Figure 5.2 would work as one, querying Twitter API right away, as soon as the tweets were consumed.

We evaluate the experiment by analyzing the metrics contained in the three distinct outputs (CSV files) from phase 3:

- **Account**
- **Tweets**
- **IoCs**

6.1 Account

The first execution has been performed and calculated with Equation 4.8, i.e. the default weighted average for the dimensions [D1..D7]. The result is represented in Table 6.2, with all the 54 accounts.
The result is represented in Figure 6.1 and shows us low account score values (HWAS - Homogeneous Weights Account Score), between 0.2388 and 0.7153, that are not representative of the real value of these accounts and its tweets. These AccountScore's are a consequence of the use of default weight values in Equation 4.8, i.e. each dimension [D1..D7] has 1/7 of weight.

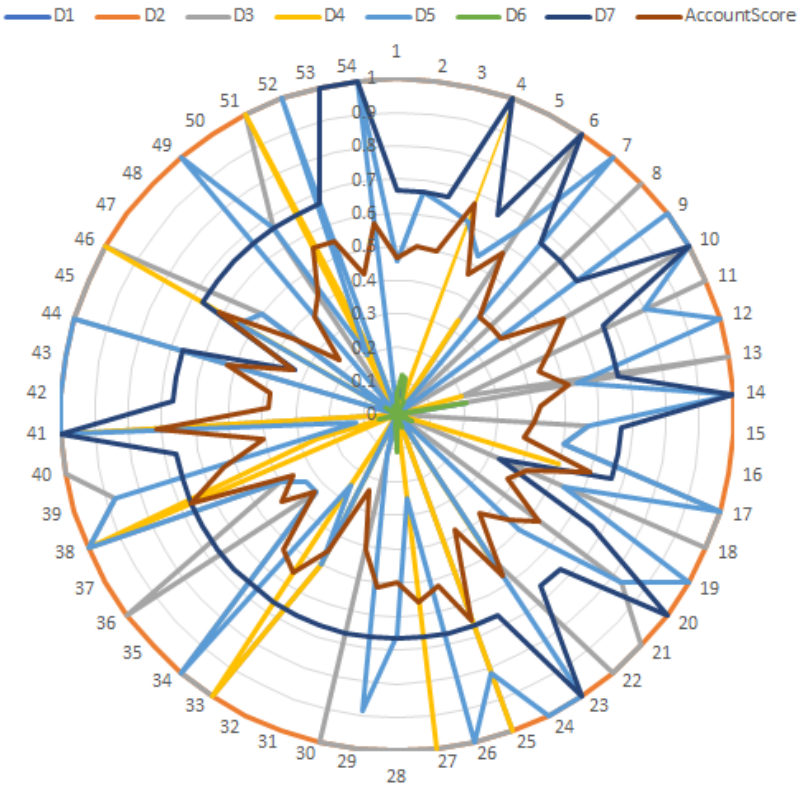


Figure 6.1: Accounts with default weights (first execution)

Figure 6.2: Account @*CVEnew* on Twitter

As example, @*CVEnew* (Figure 6.2), the account that most contributed to this dataset, has its *AccountScore* calculated in **0.51807123**. It does not reflect its true value, mainly because of D4 and D6 low values - see number 13 in Figure 6.1, @*CVEnew* has:

- **101** consumed tweets
- **101** accurate tweets (100% of its tweets)
- **101** verifiable tweets (100% of its tweets)
- **0** intelligent tweets (0% of its tweets)
- **55** timely tweets (54% of its tweets)
- **58** participated clusters (21% of the system)
- Global score **51%**

Customizing those values allow us to adapt the Equation 4.8 result according to the analysts intuition, and to what best fits the monitoring of the defined infrastructure.

D	Metric name	Weight	New weight	Justification
1	Maintenance	1/7	1/10	Not so maintained accounts shouldn't be despised
2	Accuracy	1/7	1/4	To privilege accurate accounts
3	Verifiability	1/7	1/4	To privilege accounts that provide sources
4	Intelligence	1/7	1/20	Bot-fed accounts usually do not provide external objects
5	Timeliness	1/7	1/5	To privilege timely accounts, important to have a faster response
6	Completeness	1/7	1/20	Some accounts focus on a singly type of threats or IoCs
7	Profile	1/7	1/10	Not that easy to have an account verified, even for reputable institutions

Table 6.1: Custom weights for Equation 4.8

With the custom weight values defined in Table 6.1, we have recalculated all the account scores and obtained the column *CWAS (Custom Weights AccountScore)* in Table 6.2, with scores more close to the real value of the analysed accounts, as Figure 6.3 shows. This weight customization is proposed to be allowed to the end-user through the DiSIEM OTD web interface, as covered in Figure 5.11.

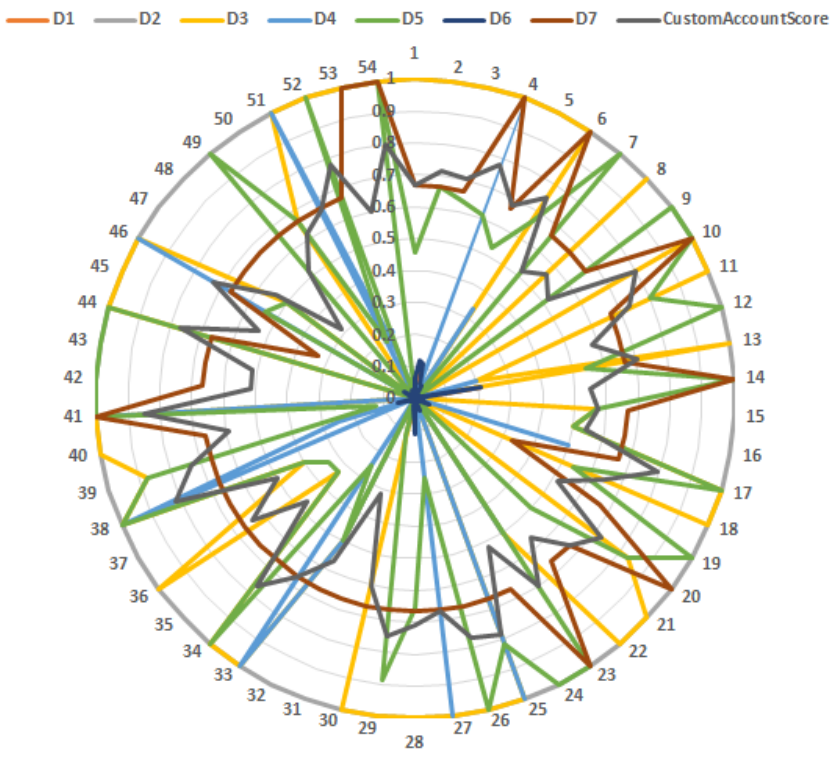


Figure 6.3: Accounts with custom weights (second execution)

#	Account	D1	M	D2	A	D3	V	D4	I	D5	T	D6	C	D7	HWAS	CWAS
1	Dinosn	0.045	22	1	22	1	22	0.045	1	0.45	10	0.0645	18	0.66	0.46805105	0.66759217
2	IT_securitynews	0.08	39	1	39	1	39	0	0	0.667	26	0.1219	34	0.66	0.50503993	0.71410167
3	threatmeter	0.09	44	1	44	1	44	0	0	0.64	28	0.11	31	0.66	0.5006415	0.70852917
4	helpnetsecurity	0.037	18	1	18	1	18	1	18	0.61	11	0.0466	13	1	0.67066675	0.7782475
5	SecurityWeek	0.035	17	1	17	1	17	0	0	0.53	9	0.0466	13	0.66	0.4682259	0.67836917
6	threatintel	0.006	3	1	3	1	3	0.33	1	0.667	2	0.0107	3	1	0.57384473	0.751154
7	MicroFocusSec	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
8	malware_traffic	0.002	1	1	1	1	1	0	0	0	0	0.0036	1	0.66	0.38175774	0.56705067
9	CyberWarship	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
10	CERTEU	0.002	1	1	1	1	1	0	0	1	1	0.0036	1	1	0.57223386	0.800384
11	InfosecurityMag	0.01	5	1	5	1	5	0	0	0.8	4	0.0143	4	0.66	0.4987529	0.72841067
12	kmkz_security	0.01	5	1	5	0.2	1	0.2	1	1	5	0.0179	5	0.66	0.44212213	0.57858967
13	CVEnew	0.207	101	1	101	1	101	0	0	0.544	55	0.2079	58	0.66	0.51807123	0.70671017
14	gcluley	0.002	1	1	1	0	0	0	0	1	1	0.0037	1	1	0.42937678	0.550384
15	cyb3rops	0.014	7	1	7	0.57	4	0	0	0.57	4	0.0251	7	0.66	0.40699816	0.57650167
16	hackerfantastic	0.008	4	1	4	0.5	2	0	0	0.5	2	0.0143	4	0.66	0.38417387	0.54320467
17	binitamshah	0.008	4	1	4	1	4	0.5	2	1	4	0.0143	4	0.66	0.5984596	0.79320467
18	MaldicoreAlerts	0.027	13	1	13	1	13	0	0	0.538	7	0.0466	13	0.33	0.42072627	0.646023834
19	shodanhq	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
20	TheHackersNews	0.012	6	1	6	0.83	5	0	0	0.83	5	0.0179	5	1	0.5281297	0.7271265
21	linuxtoday	0.008	4	1	4	1	4	0	0	0.5	2	0.0143	4	0.66	0.45560244	0.66820467
22	YoKoAcc	0.002	1	1	1	1	1	0	0	0	0	0.0036	1	0.66	0.38175774	0.56705067
23	GossiTheDog	0.004	2	1	2	0.5	1	0.5	1	1	2	0.0072	2	1	0.57303935	0.7007695
24	JAMESWT_MHT	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
25	ptracesecurity	0.0225	11	1	11	1	11	1	11	0.818	9	0.0394	11	0.66	0.64955175	0.78453317
26	x0rz	0.002	1	1	1	1	1	0	0	1	1	0.0036	1	0.66	0.5246149	0.76705067
27	KitPloit	0.008	4	1	4	1	4	1	4	0.25	1	0.0107	3	0.66	0.56223327	0.66802517
28	SecurityNewsbot	0.066	32	1	32	1	32	0	0	0.656	21	0.11	31	0.66	0.49996233	0.71004317
29	Sec_Cyber	0.018	9	1	9	1	9	0	0	0.889	8	0.0323	9	0.66	0.5151849	0.74790567
30	packet_storm	0.084	41	1	41	1	41	0	0	0.122	5	0.0251	7	0.66	0.4139852	0.60073017
31	VK_Intel	0.002	1	1	1	0	0	0	0	0	0	0.0036	1	0.66	0.23890063	0.31705067
32	DMBisson	0.004	2	1	2	0.5	1	0.5	1	0.5	1	0.0072	2	0.66	0.4539917	0.56743617
33	inj3ct0r	0.008	4	1	4	1	4	1	4	0.25	1	0.0036	1	0.66	0.5612092	0.66766667
34	424f424f	0.002	1	1	1	1	1	0	0	1	1	0.0036	1	0.66	0.5246149	0.76705067
35	thegrugq	0.006	3	1	3	0.33	1	0	0	0.33	1	0.0107	3	0.66	0.33574945	0.46781867
36	SearchSecurity	0.006	3	1	3	1	3	0	0	0.33	1	0.0072	2	0.66	0.43047553	0.63430717
37	NytroRST	0.01	5	1	5	0.4	2	0	0	0.4	2	0.0179	5	0.66	0.35640782	0.49858967
38	sans_isc	0.004	2	1	2	1	2	1	2	1	2	0.0072	2	0.66	0.6682773	0.81743617
39	VulmonFeeds	0.033	16	1	16	0.875	14	0.25	4	0.875	14	0.0538	15	0.66	0.5361835	0.72888967
40	securityaffairs	0.016	8	1	8	1	8	0	0	0.125	1	0.0143	4	0.66	0.40320438	0.59402667
41	NCSC	0.004	2	1	2	1	2	1	2	1	2	0.0072	2	1	0.7158964	0.8507695
42	circl_lu	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
43	PyroTek3	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
44	dangoodin001	0.002	1	1	1	1	1	0	0	1	1	0.0036	1	0.66	0.5246149	0.76705067
45	_aliardic_	0.002	1	1	1	1	1	0	0	0	0	0.0036	1	0.33	0.33413866	0.533717334
46	EHackerNews	0.0226	11	1	11	1	11	1	11	0.545	6	0.0394	11	0.66	0.61059076	0.72998717
47	n00py1	0.004	2	1	2	0.5	1	0	0	0.5	1	0.0072	2	0.66	0.38256314	0.54243617
48	James_inthe_box	0.002	1	1	1	0	0	0	0	0	0	0.0036	1	0.66	0.23890063	0.31705067
49	domchell	0.002	1	1	1	0	0	0	0	1	1	0.0036	1	0.66	0.38175774	0.51705067
50	byt3bl33d3r	0.006	3	1	3	0.67	2	0	0	0.66	2	0.0107	3	0.66	0.43098757	0.61782167
51	LinuxSec	0.01	5	1	5	1	5	1	5	0.2	1	0.0143	4	0.66	0.55589575	0.65841067
52	the_yellow_fall	0.0144	7	1	7	1	7	0.143	1	1	7	0.0251	7	0.66	0.5498553	0.77650117
53	HackRead	0.002	1	1	1	1	1	0	0	0	0	0.0036	1	1	0.42937678	0.600384
54	slashdot	0.01	5	1	5	1	5	0	0	1	5	0.0179	5	1	0.57545537	0.801923

Table 6.2: Account Scores with default weights and custom weights (defined in Table 6.1)

While the *AccountScore* with the default weights *(HWAS)* scored 51% for the account *@CVEnew* (#13), *CWAS* scores 70.6%, more inline with the real value of the account, with all the tweets being accurate and verifiable. What weighs negatively on the assessment of this account is D5 (Timeliness), with 0% of the tweets being timely.

6.2 Tweets

The first execution has been performed and calculated with Equation 4.14, i.e. the default weighted average for the dimensions [D8..D12]. A sample from the result is represented in Table 6.4, containing the best *TweetScore* (renamed to HWTS - Homogeneous Weights Tweet Score) values, grouped by account.

From Figure 6.4, we can observe that *BinaryConfidenceMCD* is much more optimistic about the likelihood of containing and not containing valuable information, while *TweetScore* is humble when scoring tweets, i.e., we can observe blue polygons (representing *TweetScore*) between 0.2 and 0.9, while the orange square (representing *BinaryConfidenceMCD*) are present across the whole Y axis. But, as it is reported with accounts (Chapter 6.1), *TweetScore* computation (based on the Equation 4.14) and its weights can also be customized.

Note that ”Column 1 (Tweet ID)” and ”Column 2 (Tweet)” are not represented in Table 6.4 in order to have enough room to present all the dimensions in a single page, with a readable font size.

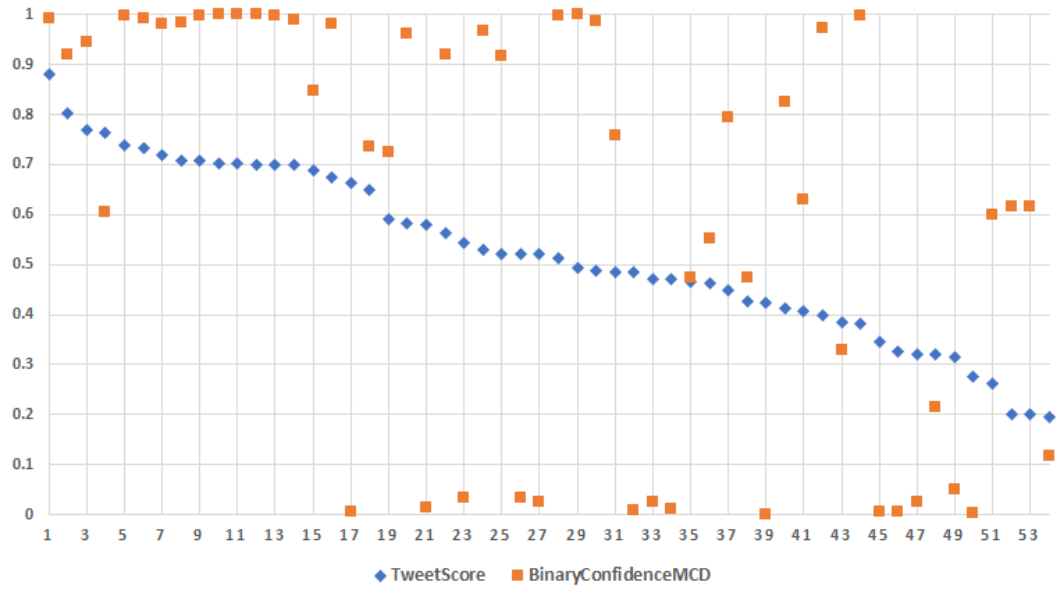


Figure 6.4: TweetScore vs BinaryConfidenceMCD (having TweetScore ordered from max. to min., from Table 6.4)

With the custom weight values defined in Table 6.3, we have recalculated all the tweet scores and obtained the column *CWTS (Custom weights TweetScore)* in Table 6.3. This weight customization is proposed to be allowed to the end-user through the DiSIEM OTD web interface, as covered in Figure 5.11.

D	Metric name	Weight	New weight	Justification
8	Validity	1/5	11/20	Above all, the impost important metric
9	Relevance	1/5	3/20	Irrelevant for DiSIEM OTD, as no weights can defined per asset
10	Timeliness	1/5	3/20	Can be somehow undervalued as it is measured in D11
11	Source account	1/5	1/10	DiSIEM OTD already focus on a restrict set of accounts
12	Tweet features	1/5	1/20	To not harm recent tweets, as the public metrics are all 0 when a tweet is posted

Table 6.3: Custom weights for Equation 4.14

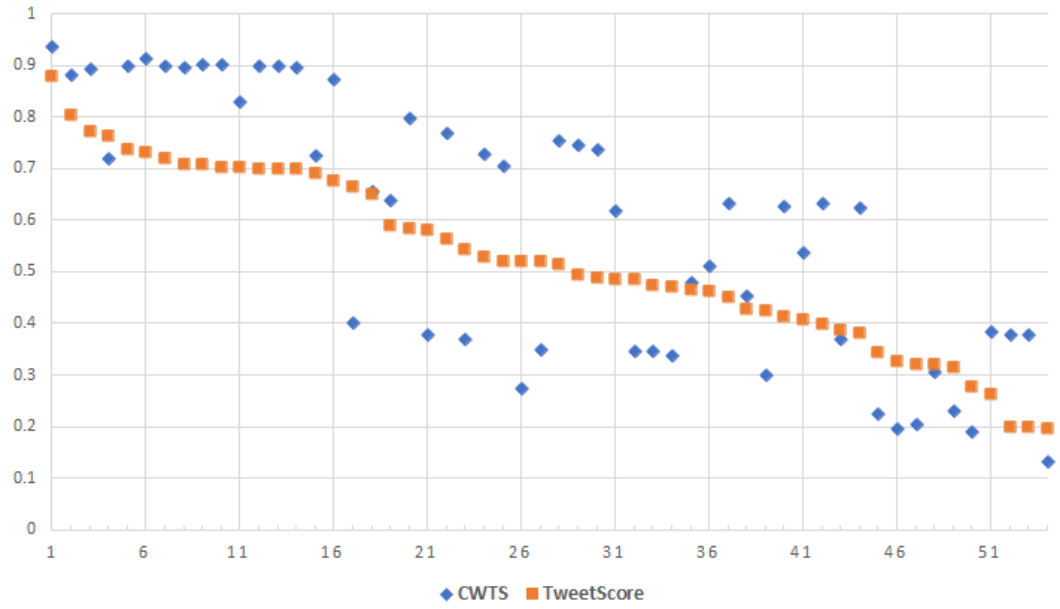


Figure 6.5: TweetScore vs CWTS (having TweetScore ordered from max. to min., from Table 6.4)

As Figure 6.5 shows, *CWTS* is not always more optimistic than *TweetScore*, as we have some blue polygons below the orange squares. It outputs the real value of the information according to the defined metrics weights.

Account	D8	D9	D10	D11	D12	HWTS	BMCD	CWTS
cyb3rops	0.99168915	1	1	0.406998158	1	0.879737462	0.99168915	0.936129
Dinosn	0.91977197	1	0.965	0.468051046	0.66	0.803897937	0.91977197	0.880763
slashdot	0.94433415	1	1	0.575455368	0.33	0.77062457	0.94433415	0.893596
inj3ct0r	0.6061339	1	0.98	0.561209202	0.66	0.762801954	0.6061339	0.719828
circl_lu	0.9965403	1	0.975	0.381757736	0.33	0.737326274	0.9965403	0.89919
helpnetsecurity	0.99125516	1	1	0.670666754	0	0.732384383	0.99125516	0.912257
ptracesecurity	0.98064274	1	0.965	0.649551749	0	0.719038898	0.98064274	0.899059
LinuxSec	0.98402	1	1	0.555895746	0	0.707983149	0.98402	0.896801
VulmonFeeds	0.9989229	1	1	0.536183476	0	0.707021275	0.9989229	0.903026
CVEnew	0.9993175	1	1	0.518071234	0	0.703477747	0.9993175	0.901432
EHackerNews	0.9996207	1	0.24	0.610590756	0.66	0.703375625	0.9996207	0.830184
threatmeter	0.9993728	1	1	0.500641525	0	0.700002865	0.9993728	0.899719
SecurityNewsbot	0.99867046	1	1	0.49996233	0	0.699726558	0.99867046	0.899265
IT_securitynews	0.9901444	1	1	0.50503993	0	0.699036866	0.9901444	0.895083
binitamshah	0.84815395	1	0	0.598459601	1	0.68932271	0.84815395	0.726331
Sec_Cyber	0.9801033	1	0.88	0.515184879	0	0.675057636	0.9801033	0.872575
sans_isc	0.006643932	1	0.975	0.668277323	0.66	0.663317584	0.006643932	0.400065
TheHackersNews	0.73718184	0	0.98	0.528129697	1	0.649062307	0.73718184	0.655263
KitPloit	0.7246598	0	1	0.562233269	0.66	0.590711947	0.7246598	0.63812
InfosecurityMag	0.96144116	1	0.455	0.498752892	0	0.58303881	0.96144116	0.796918
the_yellow_fall	0.013678286	1	1	0.549855292	0.33	0.579373382	0.013678286	0.379175
MaldicoreAlerts	0.91977197	1	0.475	0.420726269	0	0.563099648	0.91977197	0.769197
malware_traffic	0.03301796	1	0.97	0.381757736	0.33	0.543621806	0.03301796	0.368502
kmkz_security	0.9667935	0	0.905	0.442122132	0.33	0.529449793	0.9667935	0.728365
NytroRST	0.91695184	1	0	0.356407821	0.33	0.521338599	0.91695184	0.706631
GossiTheDog	0.032692973	0	1	0.573039353	1	0.521146465	0.032692973	0.275285
hackerfantastic	0.025396591	1	0.86	0.38417387	0.33	0.520580759	0.025396591	0.348052
threatintel	0.997106	1	0	0.573844731	0	0.514190146	0.997106	0.755793
SecurityWeek	0.9999478	0	1	0.468225896	0	0.493634739	0.9999478	0.746794
linuxtoday	0.9860508	1	0	0.455602437	0	0.488330647	0.9860508	0.737888
aliardic	0.7581368	1	0	0.334138662	0.33	0.485121759	0.7581368	0.617056
packet_storm	0.007673061	1	1	0.413985193	0	0.484331651	0.007673061	0.345619
thegrugq	0.025753044	1	1	0.335749447	0	0.472300498	0.025753044	0.347739
byt3bl33d3r	0.012099793	1	0.915	0.430987567	0	0.471617472	0.012099793	0.337004
x0rz	0.47370282	0	0.99	0.524614871	0.33	0.464330205	0.47370282	0.478165
HackRead	0.5506387	1	0	0.429376781	0.33	0.462669763	0.5506387	0.512456
DMBisson	0.7955018	1	0	0.453991711	0	0.449898702	0.7955018	0.632925
SearchSecurity	0.47370282	0	0.895	0.430475533	0.33	0.426502337	0.47370282	0.454501
MicroFocusSec	0.001075793	1	0.74	0.381757736	0	0.424566706	0.001075793	0.299767
James_inthe_box	0.8256198	1	0	0.238900632	0	0.412904086	0.8256198	0.627981
securityaffairs	0.62858117	1	0	0.403204381	0	0.40635711	0.62858117	0.53604
gcluley	0.9732938	0	0.255	0.429376781	0.33	0.398200783	0.9732938	0.633166
PyroTek3	0.32932225	0	0.885	0.381757736	0.33	0.385882664	0.32932225	0.36872
CERTEU	0.9982147	0	0	0.572233856	0.33	0.380756378	0.9982147	0.622908
NCSC	0.007312276	1	0	0.715896428	0	0.344641741	0.007312276	0.225611
domchell	0.006140062	0	0.915	0.381757736	0.33	0.327246226	0.006140062	0.195469
VK_Intel	0.025396591	1	0	0.238900632	0.33	0.319526111	0.025396591	0.204525
JAMESWT_MHT	0.21571183	1	0	0.381757736	0	0.319493913	0.21571183	0.306817
424f424f	0.05036847	1	0	0.524614871	0	0.314996668	0.05036847	0.230164
shodanhq	0.003116445	1	0	0.381757736	0	0.276974836	0.003116445	0.18989
YoKoAcc	0.60006297	0	0	0.381757736	0.33	0.263030808	0.60006297	0.384877
n00py1	0.61713904	0	0	0.382563144	0	0.199940437	0.61713904	0.377683
CyberWarship	0.61713904	0	0	0.381757736	0	0.199779355	0.61713904	0.377602
dangoodin001	0.11705124	0	0	0.524614871	0.33	0.194999889	0.11705124	0.133506

Table 6.4: Tweets Assessment - Best HWTS and CWTS per account (sample)

6.3 IoCs

The *IoCScore*, based on the maximum *TweetScore* within a cluster, inherit its properties, i.e. it is calculated with the Equation 4.14, and depends of the weights defined for [D8..D12] in the *TweetScore* computation.

The first execution has been performed and calculated with the Equation 4.14, i.e. the default weighted average for the dimensions [D8..D12]. The result is represented in Figure 6.6 and Table 6.5 contains in detail the 3rd output produced.

Similarly to what happens with *TweetScore*, *IoCScore* is by default very humble at classifying, when comparing to *BinaryConfidenceMCD*. The ability of customizing weights allows the final user to take advantage of it, by defining the values accordingly to his reality, and by better understanding the score, as it is not just a simple analytic that the user does not control or knows how it is calculated.

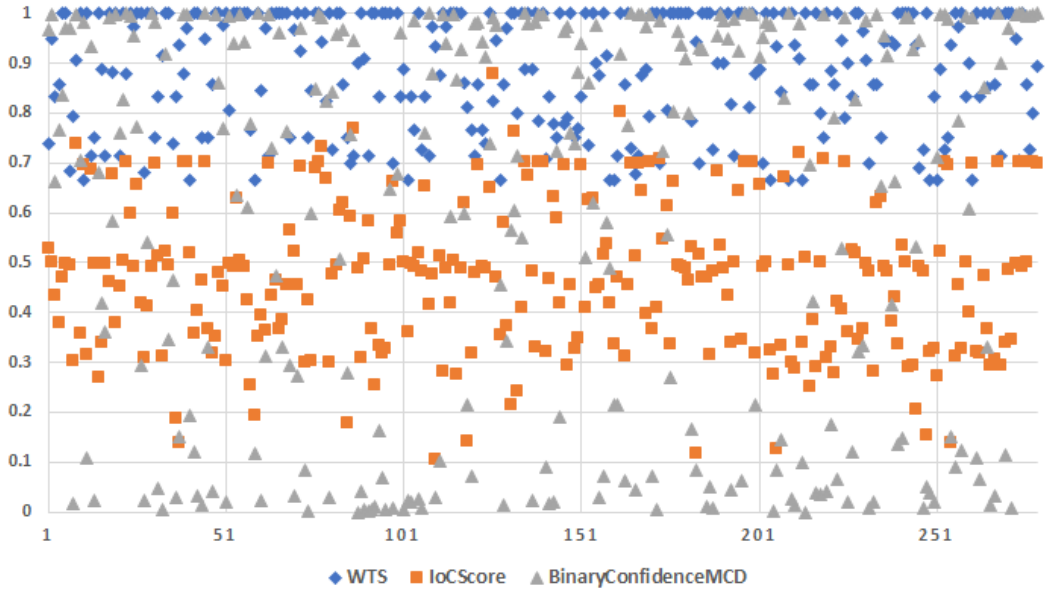


Figure 6.6: IoC analysis

The cluster size distribution (see Figure 6.6) is on average 2.74, meaning that there were very few IoCs gathered by DiSIEM OTD with tweets referencing the same threat, or at least they were not grouped due to some unknown reason. One IoC with 10 and other with 21 tweets stand out due to its size.

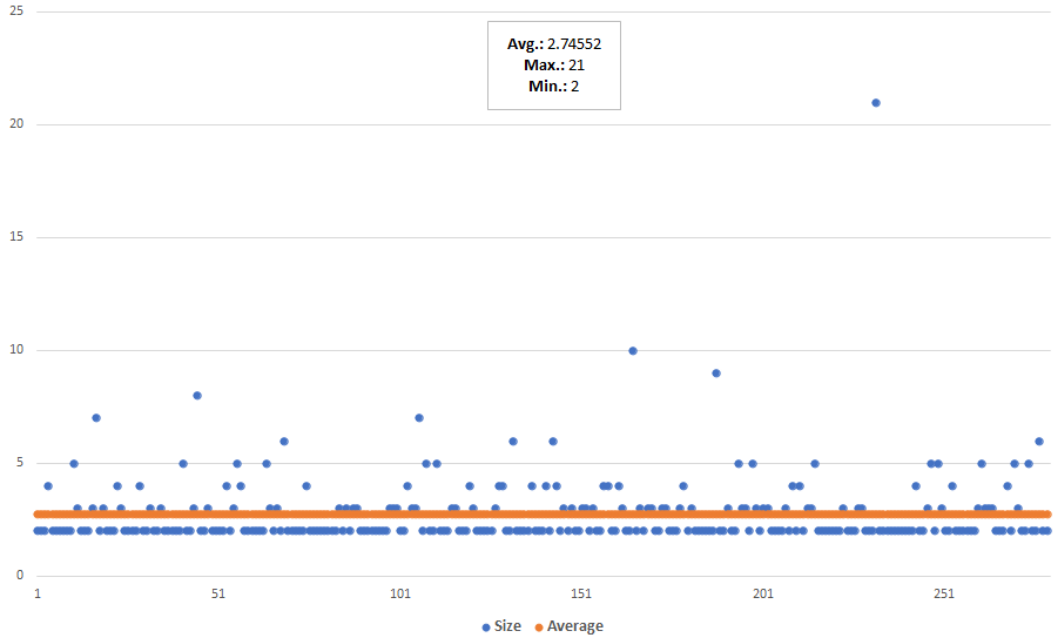


Figure 6.7: Cluster size distribution and average

The list of collected threat identifiers (see Table 6.6) counts with an average 7.64 (HIGH) CVSS3 base score (see Appendix B), meaning that the monitored tweets reference highly classified CVEs, with at least two pointing to a vulnerability later classified with CVSS 10 (CRITICAL). Some IDs from RedHat Security Advisory (RHSA) were also identified and collected also reference in some cases HIGH vulnerabilities - RHSA-2021:2663 (CVE-2021-3583), RHSA-2021:2664 (CVE-2021-3583), RHSA-2021:2693 (CVE-2021-3536, CVE-2021-21409), RHSA-2021:2694 (CVE-2021-3536, CVE-2021-21410), RHSA-2021:2716 (CVE-2021-32399, CVE-2021-33909), RHSA-2021:2720 (CVE-2021-33034, CVE-2021-33909), RHSA-2021:3381 (CVE-2021-22555, CVE-2021-32399), RHSA-2021:3392 (CVE-2021-32399), RHSA-2021:3443 (CVE-2021-0512, CVE-2021-3715, CVE-2021-37576).

When it comes to threats and keywords (see Table 6.7), the dataset was full of different threats, impacting a huge variety of assets from the defined infrastructure.

Size	ThreatList	ThreatIDs	Keywords	LinksList	LastUpdated	Tweet	IoCScore	WTS	BMCD
2	[exploit, vulnerability]	[CVE-2021-33909]	[linux]	[https://t.co/ZRtZmuGz1B]	7/21/2021 7:41	RT @cyb3rops: Exploi…	0.879737462	0.823529412	0.99168915
3	[attack]	[]	[office, windows]	[https://t.co/rUl8HpdA8V]	9/8/2021 17:00	New 0-Day Attack Tar…	0.803897937	1	0.91977197
2	[]	[]	[linux]	[https://t.co/YnirOvfJU9]	8/17/2021 1:30	Linux glibc security…	0.77062457	0.714285714	0.94433415
6	[exploit, 0day, remote, execution]	[]	[exchange]	[https://t.co/KoPTM30D9s, https://t.co/DkmuVfKI5h]	7/27/2021 6:17	#0day #Microsoft Exc…	0.762801954	1	0.6061339
2	[exploit, vulnerabilities]	[]	[chrome]	[https://t.co/NRHbH5d5TG]	7/16/2021 15:08	Google has released …	0.737326274	0.904761905	0.9965403
2	[phishing]	[]	[office]	[https://t.co/VvuTE8Q3qb]	8/8/2021 8:30	Week in review: Clev…	0.732384383	0.941176471	0.99125516
2	[exploit, exploitation]	[]	[linux]	[https://t.co/PVkZCiNcGd, https://t.co/k2Auvp5O4k]	8/18/2021 10:07	Learning Linux Kerne…	0.719038898	0.909090909	0.98064274
3	[attack, vulnerability]	[]	[solarwinds]	[https://t.co/7z97zyR2aZ]	7/13/2021 13:18	A New #Critical Sola…	0.707983149	0.7	0.98402
2	[vulnerability]	[CVE-2021-30592Out, CVE-2021-30594Use]	[chrome]	[https://t.co/3oqrWAfdsU, https://t.co/oh1APvMMnd]	8/26/2021 21:12	CVE-2021-30592Out of…	0.707021275	0.75	0.9989229
4	[attack, exploit, remote]	[CVE-2021-30569, CVE-2021-30572, CVE-2021-30573, CVE-2021-30574]	[chrome]	[https://t.co/VzDwLVr0jE]	8/5/2021 19:38	CVE-2021-30572 Use a…	0.703477747	0.878787879	0.9993175
5	[attack, remote, leak, exploit]	[CVE-2021-30617, CVE-2021-30621, CVE-2021-30619, CVE-2021-30615, ...]	[chrome]	[https://t.co/q6cZztH8zS]	9/3/2021 20:45	CVE-2021-30621 Inapp…	0.703468747	0.705882353	0.9992725
5	[attack, exploit, remote]	[CVE-2021-30559, CVE-2021-30560, CVE-2021-30561, CVE-2021-30562, ...]	[chrome]	[https://t.co/m6Gd0M9PBD]	8/5/2021 19:38	CVE-2021-30561 Type …	0.703457297	0.8125	0.99921525
8	[attack, exploit, remote, compromise]	[CVE-2021-30606, CVE-2021-30607, CVE-2021-30608, CVE-2021-30609, ...]	[chrome]	[https://t.co/DrlM8UQAi7]	9/3/2021 20:45	CVE-2021-30606 Use a…	0.703456227	0.95	0.9992099
4	[attack, malicious, exploit, remote, compromise]	[CVE-2021-30601, CVE-2021-30600, CVE-2021-30603, CVE-2021-30604]	[chrome]	[https://t.co/kGdL6y7U33]	8/26/2021 18:45	CVE-2021-30601 Use a…	0.703453427	1	0.9991959
2	[attack, remote, exploit]	[CVE-2021-30589, CVE-2021-30588]	[chrome]	[https://t.co/5cQXTlkECP]	8/5/2021 19:38	CVE-2021-30589 Insuf…	0.703423787	1	0.9990477
4	[attack, remote, access, escalation, malicious, exploit]	[CVE-2021-30578, CVE-2021-30577, CVE-2021-30576, CVE-2021-30579]	[chrome]	[https://t.co/82FNunw205]	8/5/2021 19:38	CVE-2021-30578 Unini…	0.703402547	0.727272727	0.9989415
2	[exploit]	[]	[linux]	[https://t.co/PJevS5cceY]	8/1/2021 18:38	Exploit Code Release…	0.703375625	0.888888889	0.9996207
2	[attack, malicious, exploit, remote]	[CVE-2021-30565, CVE-2021-30568]	[chrome, linux]	[https://t.co/B4yVdwZxY2]	8/5/2021 19:38	CVE-2021-30565 Out o…	0.703304427	1	0.9984509
2	[attack, remote]	[CVE-2021-30598, CVE-2021-30599]	[chrome]	[https://t.co/ZVoU0p28Ux]	8/26/2021 18:45	CVE-2021-30599 Type …	0.703299707	0.970588235	0.9984273
2	[attack, remote]	[CVE-2021-30620]	[chrome]	[https://t.co/mNere17ZNM]	9/3/2021 20:45	CVE-2021-30620 Insuf…	0.703283655	0.8	0.99834704
2	[attack, exploit, remote, access]	[CVE-2021-30594, CVE-2021-30597]	[chrome]	[https://t.co/22DZLHb0OC]	8/26/2021 17:45	CVE-2021-30597 Use a…	0.703281867	0.878787879	0.9983381
3	[attack, spoofing, remote, leak]	[CVE-2021-30584, CVE-2021-30582, CVE-2021-30583]	[chrome]	[https://t.co/SG2aOze3fa]	8/5/2021 19:38	CVE-2021-30584 Incor…	0.703264057	1	0.99824905
2	[attack, exploit, remote, compromise]	[CVE-2021-30566, CVE-2021-30564]	[chrome]	[https://t.co/hvAX2N7Yy9]	8/5/2021 19:38	CVE-2021-30566 Stack…	0.703197193	1	0.99791473
3	[attack, malicious, exploit]	[CVE-2021-30622, CVE-2021-30623]	[chrome]	[https://t.co/tirxsWmRNZ]	9/4/2021 7:09	CVE-2021-30622 Use a…	0.703088467	0.794871795	0.9973711
2	[attack, malicious, exploit]	[CVE-2021-30580, CVE-2021-30581]	[chrome]	[https://t.co/wokeAYcDX8]	8/5/2021 19:38	CVE-2021-30581 Use a…	0.703012387	0.783783784	0.9969907
6	[attack, malicious, exploit]	[CVE-2021-30611, CVE-2021-30612, CVE-2021-30614]	[chrome]	[https://t.co/9rSwRMuSO1]	9/4/2021 7:09	CVE-2021-30612 Use a…	0.702647997	0.725	0.99516875
2	[]	[CVE-2021-40490]	[linux]	[https://t.co/HYE9zl9B9K]	9/3/2021 1:45	CVE-2021-40490 A rac…	0.702037407	0.791666667	0.9921158
2	[denial of service, vulnerability]	[CVE-2021-1112, CVE-2021-1114]	[linux]	[https://t.co/piJIcoGbgM]	8/11/2021 22:45	CVE-2021-1112 NVIDIA…	0.700041973	0.75	0.98213863
2	[attack, malicious, exploit, remote]	[CVE-2021-30623, CVE-2021-30624]	[chrome]	[https://t.co/P5aMQa82s2]	9/4/2021 7:09	CVE-2021-30624 Use a…	0.700002865	0.894736842	0.9993728
10	[attack, remote, leak, exploit, compromise]	[CVE-2021-30617, CVE-2021-30621, CVE-2021-30619, CVE-2021-30615, ...]	[chrome]	[https://t.co/5MM1uCIdef]	9/4/2021 7:09	CVE-2021-30621 Inapp…	0.699991239	0.72972973	0.99931467
2	[]	[(RHSA-2021:2694), (RHSA-2021:2693)]	[rhel]	[https://t.co/T9pxDibDxh]	7/17/2021 9:00	RHEL 8 : Red Hat JBo…	0.699726558	0.714285714	0.99867046
3	[]	[(RHSA-2021:3381), (RHSA-2021:3392), (RHSA-2021:3443)]	[rhel]	[https://t.co/h2ANiJaiQ2]	9/5/2021 1:00	RHEL 7 : kpatch-patc…	0.699229726	0.714285714	0.9961863
5	[]	[]	[sap]	[https://t.co/B69m68Saz4]	8/11/2021 16:38	SAP patches critical…	0.699036866	0.666666667	0.9901444
2	[vulnerability]	[]	[linux]	[https://t.co/48cnJ3GuB6]	8/2/2021 11:50	Linux Kernel up to 5…	0.697573555	0.75	0.98723525
2	[exploit, vulnerability]	[]	[chrome]	[https://t.co/q95cT9hby1, https://t.co/F2zbXmBs16]	7/17/2021 10:44	RT @CyberSecurityPr…	0.697274726	0.857142857	0.9813337
5	[vulnerability]	[]	[chrome]	[https://t.co/rBloRYL4EP]	7/18/2021 7:43	Google Chrome prior …	0.696787655	0.666666667	0.98329675
3	[attack, vulnerability, remote]	[CVE-2018-17861, CVE-2018-17862, CVE-2018-17865]	[sap]	[https://t.co/71AvT9R2hf]	8/9/2021 19:45	CVE-2018-17861 ** UN…	0.696506827	0.777777778	0.9644629
3	[vulnerability, denial of service, execution, access]	[CVE-2021-1106, CVE-2021-1112, CVE-2021-1107]	[linux]	[https://t.co/1j2xZmkOkX, https://t.co/xhLrNnfyrO]	8/12/2021 1:23	CVE-2021-1106NVIDIA …	0.695284215	0.833333333	0.9402376
2	[]	[(RHSA-2021:2716), (RHSA-2021:2720)]	[rhel]	[https://t.co/WW2M9foPAO]	7/26/2021 17:15	RHEL 8 : kpatch-patc…	0.694217718	0.923076923	0.97112626
2	[]	[]	[linux]	[https://t.co/meIvqdwSAT]	8/11/2021 5:52	The Linux Kernel Mod…	0.68932271	1	0.84815395
2	[injection]	[]	[sap]	[https://t.co/qypgXdKJ6j]	7/21/2021 7:00	SAP NetWeaver AS ABA…	0.686931166	0.714285714	0.9346935
2	[attack]	[]	[linux]	[https://t.co/L2yw2ae4Sy]	7/21/2021 5:58	LPE flaw in Linux ke…	0.685222986	0.9	0.956075
3	[fake, phishing]	[CVE-2021-30554]	[chrome]	[https://t.co/fFo1weJWCD]	7/22/2021 14:18	HTA Files Distribute…	0.676754104	0.882352941	0.58238614
2	[phishing]	[]	[office]	[https://t.co/U6BW9rXmV9]	8/5/2021 3:42	A clever phishing ca…	0.675057636	1	0.9801033
2	[attack, remote]	[CVE-2016-1998HPE, CVE-2016-1985HPE]	[java, windows]	[https://t.co/0FisWsyQAl]	8/12/2021 16:41	CVE-2016-1998HPE Ser…	0.672954703	1	0.82859004
2	[attack, remote]	[CVE-2015-2073, CVE-2015-2074]	[sap]	[https://t.co/sYVurPZMXE]	8/9/2021 18:45	CVE-2015-2073 The Fi…	0.668596407	0.823529412	0.8249108
2	[vulnerability, remote, execution]	[CVE-2021-34518, CVE-2021-34501., CVE-2021-34501, CVE-2021-34518]	[excel]	[https://t.co/UsTBDutMhV]	7/14/2021 18:46	CVE-2021-34518 Micro…	0.663856427	1	0.8012109
3	[malware]	[]	[java]	[https://t.co/AmWIkjhqgO]	9/1/2021 0:20	RT @sans_isc: ISC di…	0.663317584	0.7	0.006643932
2	[exploit, exploitation, denial of service, attack]	[, CVE-2021-38203]	[linux]	[https://t.co/NzQTyR9b1U]	8/8/2021 20:45	Learning Linux Kerne…	0.657788627	1	0.7708719
3	[malicious, exploit]	[]	[chrome]	[https://t.co/uhpBasqCoj]	8/5/2021 3:28	Google Patches Sever…	0.65763942	0.888888889	0.91315717

Table 6.5: IoCs top 30 sample

CVE	CVSS3 Base Score	CVE	CVSS3 Base Score	CVE	CVSS3 Base Score	CVE	CVSS3 Base Score
CVE-2015-2073	7.5 - HIGH	CVE-2021-30580	6.5 - MEDIUM	CVE-2021-33678	6.5 - MEDIUM	CVE-2021-34499	6.5 - MEDIUM
CVE-2015-2074	7.5 - HIGH	CVE-2021-30581	8.8 - HIGH	CVE-2021-33680	6.5 - MEDIUM	CVE-2021-34501	8.8 - HIGH
CVE-2016-1985	10 - CRITICAL	CVE-2021-30582	6.5 - MEDIUM	CVE-2021-33681	6.5 - MEDIUM	CVE-2021-34503	7.8 - HIGH
CVE-2016-1998	9.8 - CRITICAL	CVE-2021-30583	6.5 - MEDIUM	CVE-2021-33682	5.4 - MEDIUM	CVE-2021-34508	8.8 - HIGH
CVE-2017-11176	7.8 - HIGH	CVE-2021-30584	6.5 - MEDIUM	CVE-2021-33687	4.9 - MEDIUM	CVE-2021-34514	7.8 - HIGH
CVE-2018-17861	6.1 - MEDIUM	CVE-2021-30588	8.8 - HIGH	CVE-2021-33689	4.3 - MEDIUM	CVE-2021-34518	7.8 - HIGH
CVE-2018-17862	6.1 - MEDIUM	CVE-2021-30589	4.3 - MEDIUM	CVE-2021-33745	6.5 - MEDIUM	CVE-2021-34523	9.8 - CRITICAL
CVE-2018-17865	6.1 - MEDIUM	CVE-2021-30592	8.8 - HIGH	CVE-2021-33746	8.8 - HIGH	CVE-2021-34525	8.8 - HIGH
CVE-2018-8453	7.8 - HIGH	CVE-2021-30594	6.8 - MEDIUM	CVE-2021-33749	8.8 - HIGH	CVE-2021-34530	7.8 - HIGH
CVE-2018-8639	7.8 - HIGH	CVE-2021-30597	6.8 - MEDIUM	CVE-2021-33750	8.8 - HIGH	CVE-2021-34533	7.8 - HIGH
CVE-2019-0836	7.8 - HIGH	CVE-2021-30598	8.8 - HIGH	CVE-2021-33752	8.8 - HIGH	CVE-2021-34687	5.3 - MEDIUM
CVE-2019-0841	7.8 - HIGH	CVE-2021-30599	8.8 - HIGH	CVE-2021-33754	8 - HIGH	CVE-2021-34688	3.3 - LOW
CVE-2019-1064	7.8 - HIGH	CVE-2021-30600	8.8 - HIGH	CVE-2021-33755	8.6 - HIGH	CVE-2021-34689	5.5 - MEDIUM
CVE-2019-1129	7.8 - HIGH	CVE-2021-30601	8.8 - HIGH	CVE-2021-33756	8.8 - HIGH	CVE-2021-34690	9.8 - CRITICAL
CVE-2020-0668	7.8 - HIGH	CVE-2021-30603	7.5 - HIGH	CVE-2021-33758	7.7 - HIGH	CVE-2021-34691	7.5 - HIGH
CVE-2020-14999	7.5 - HIGH	CVE-2021-30604	8.8 - HIGH	CVE-2021-33761	7.8 - HIGH	CVE-2021-34692	7.8 - HIGH
CVE-2021-0084	7.8 - HIGH	CVE-2021-30606	8.8 - HIGH	CVE-2021-33763	5.5 - MEDIUM	CVE-2021-35211	10 - CRITICAL
CVE-2021-1090	7.1 - HIGH	CVE-2021-30607	8.8 - HIGH	CVE-2021-33764	5.9 - MEDIUM	CVE-2021-3573	6.4 - MEDIUM
CVE-2021-1094	6.1 - MEDIUM	CVE-2021-30608	8.8 - HIGH	CVE-2021-33765	5.5 - MEDIUM	CVE-2021-3612	7.8 - HIGH
CVE-2021-1095	5.5 - MEDIUM	CVE-2021-30609	8.8 - HIGH	CVE-2021-33768	8 - HIGH	CVE-2021-36926	7.5 - HIGH
CVE-2021-1096	5.5 - MEDIUM	CVE-2021-30610	8.8 - HIGH	CVE-2021-33771	7.8 - HIGH	CVE-2021-36932	7.5 - HIGH
CVE-2021-1106	7.8 - HIGH	CVE-2021-30611	8.8 - HIGH	CVE-2021-33772	7.5 - HIGH	CVE-2021-36933	7.5 - HIGH
CVE-2021-1107	7.8 - HIGH	CVE-2021-30612	8.8 - HIGH	CVE-2021-33773	7.8 - HIGH	CVE-2021-36934	7.8 - HIGH
CVE-2021-1108	7.3 - HIGH	CVE-2021-30613	8.8 - HIGH	CVE-2021-33779	6.5 - MEDIUM	CVE-2021-36936	9.8 - CRITICAL
CVE-2021-1112	5.5 - MEDIUM	CVE-2021-30614	8.8 - HIGH	CVE-2021-33780	8.8 - HIGH	CVE-2021-36945	7.8 - HIGH
CVE-2021-1114	4.4 - MEDIUM	CVE-2021-30615	6.5 - MEDIUM	CVE-2021-33782	5.5 - MEDIUM	CVE-2021-36947	8.8 - HIGH
CVE-2021-1955	7.5 - HIGH	CVE-2021-30616	8.8 - HIGH	CVE-2021-33784	7.8 - HIGH	CVE-2021-36948	7.8 - HIGH
CVE-2021-22921	7.8 - HIGH	CVE-2021-30617	6.5 - MEDIUM	CVE-2021-33909	7.8 - HIGH	CVE-2021-36958	7.8 - HIGH
CVE-2021-36948	7.8 - HIGH	CVE-2021-30618	8.8 - HIGH	CVE-2021-34438	7.8 - HIGH	CVE-2021-37573	6.1 - MEDIUM
CVE-2021-26425	7.8 - HIGH	CVE-2021-30619	6.5 - MEDIUM	CVE-2021-34439	7.8 - HIGH	CVE-2021-37576	7.8 - HIGH
CVE-2021-26433	7.5 - HIGH	CVE-2021-30620	8.8 - HIGH	CVE-2021-34441	7.8 - HIGH	CVE-2021-37628	7.5 - HIGH
CVE-2021-30554	8.8 - HIGH	CVE-2021-30621	6.5 - MEDIUM	CVE-2021-34442	7.5 - HIGH	CVE-2021-37629	5.3 - MEDIUM
CVE-2021-30559	8.8 - HIGH	CVE-2021-30622	8.8 - HIGH	CVE-2021-34444	6.5 - MEDIUM	CVE-2021-38086	7.8 - HIGH
CVE-2021-30560	8.8 - HIGH	CVE-2021-30623	8.8 - HIGH	CVE-2021-34445	7.8 - HIGH	CVE-2021-38088	7.8 - HIGH
CVE-2021-30561	8.8 - HIGH	CVE-2021-30624	8.8 - HIGH	CVE-2021-34454	5.5 - MEDIUM	CVE-2021-38166	7.8 - HIGH
CVE-2021-30562	8.8 - HIGH	CVE-2021-31183	7.5 - HIGH	CVE-2021-34456	7.8 - HIGH	CVE-2021-38203	5.5 - MEDIUM
CVE-2021-30563	8.8 - HIGH	CVE-2021-31196	7.2 - HIGH	CVE-2021-34457	5.5 - MEDIUM	CVE-2021-38208	5.5 - MEDIUM
CVE-2021-30564	8.8 - HIGH	CVE-2021-31206	8 - HIGH	CVE-2021-34458	9.9 - CRITICAL	CVE-2021-39115	7.2 - HIGH
CVE-2021-30565	8.8 - HIGH	CVE-2021-31956	7.8 - HIGH	CVE-2021-34466	6.1 - MEDIUM	CVE-2021-39177	9.8 - CRITICAL
CVE-2021-30566	8.8 - HIGH	CVE-2021-31979	7.8 - HIGH	CVE-2021-34469	8.1 - HIGH	CVE-2021-40444	7.8 - HIGH
CVE-2021-30568	8.8 - HIGH	CVE-2021-32576	7.8 - HIGH	CVE-2021-34470	8 - HIGH	CVE-2021-40490	7 - HIGH
CVE-2021-30569	8.8 - HIGH	CVE-2021-32577	7.8 - HIGH	CVE-2021-34473	9.8 - CRITICAL	RHSA-2021:2663	
CVE-2021-30571	9.6 - CRITICAL	CVE-2021-32578	7.8 - HIGH	CVE-2021-34483	7.8 - HIGH	RHSA-2021:2664	
CVE-2021-30572	8.8 - HIGH	CVE-2021-32580	7.8 - HIGH	CVE-2021-34485	5.5 - MEDIUM	RHSA-2021:2693	
CVE-2021-30573	8.8 - HIGH	CVE-2021-32769	7.5 - HIGH	CVE-2021-34486	7.8 - HIGH	RHSA-2021:2694	
CVE-2021-30574	8.8 - HIGH	CVE-2021-33667	4.3 - MEDIUM	CVE-2021-34487	7.8 - HIGH	RHSA-2021:2716	
CVE-2021-30576	8.8 - HIGH	CVE-2021-33670	7.5 - HIGH	CVE-2021-34490	7.5 - HIGH	RHSA-2021:2720	
CVE-2021-30577	7.8 - HIGH	CVE-2021-33671	8.8 - HIGH	CVE-2021-34493	6.7 - MEDIUM	RHSA-2021:3381	
CVE-2021-30578	8.8 - HIGH	CVE-2021-33676	7.2 - HIGH	CVE-2021-34494	8.8 - HIGH	RHSA-2021:3392	
CVE-2021-30579	8.8 - HIGH	CVE-2021-33677	7.5 - HIGH	CVE-2021-34496	5.5 - MEDIUM	RHSA-2021:3443	

Table 6.6: Captured Threat IDs and respective CVSS

Threat
0day
access
attack
botnet
breach
compromise
denial of service
elevation
escalation
execution
exploit
exploitation
fake
hijack
hijacking
injection
integrity
leak
malicious
malware
man in the middle
phishing
ransomware
remote
spoofing
spyware
steal
stolen
targeted attacks
vulnerabilities
vulnerability

Keyword
windows
rhel
linux
chrome
sap
exchange
office
.net
flash
solarwinds
java
excel
word
mcafee
scada

Table 6.7: Collected Threats and Keywords

From the experimental evaluation, we can conclude that the assessment of cyber threats discovered by OSINT in distinct areas and its visual representation, allows to optimize the prioritization of threat analysis and the time to analyse each IoC. The learning curve to get analysts comfortable with the metric preferences is very small and take us from a almost zero-analytics platform to a enriched-IoC solution. The real advantage of making adjustable metrics available is to set account score and/or tweet/IoC score with the best reliability assessment possible. *AccountScore*, *TweetScore* and *IoCScore* (Equations 4.8, 4.14, 4.15) performed well and revealed themselves very humble with custom weights, while the default weights presented low performance by not scoring the real value of the assessed elements.

During the experimental evaluation, we have conducted interviews with Cyber Security Professionals (see Appendix D) who have already had contact with the current version of DiSIEM OTD Platform, but who are no longer using, to the detriment of commercial solutions with vendor support, that sets them free of CTI processing and that satisfies them in terms of quality, relevance and timeliness. In their opinion, our proposal of metrics would benefit the time to act and response on the threats provided by non-commercial and OSINT sources, and would ease the focus on Twitter accounts that often tweet relevant and accurate CTI. That said, and by trusting a set of accounts, the Tweet metrics that would improve the reliability of CTI would be the validity and the relevance of the data. That would be possible, as our proposal allows the customization of the metrics weights, where the analysts could set low values to the remaining metrics and high values to those ones. However, it is also their opinion that our proposal includes too many metrics, some that could be grouped by category, when referring the same topic of interest, i.e.:

- **Account volumetry:** Maintenance (D1) and Completeness (D6)
- **External information:** Verifiability (D3) and Intelligence (D4)
- **Trustworthiness:** Validity (D8) and Relevance (D9)
- **Twitter public metrics:** Profile (D7) and Tweet features (D12)

We are of the opinion that all the suggested metrics have a purpose and that by adjusting the weights, analysts would be more than comfortable at checking only the Account Score and the Tweet/IoC score.

Chapter 7

Conclusion & Future work

In this work, we propose a set of metrics to assess the trustworthiness of cyber threats discovered on Twitter, as well as an interface mock-up of a potential integration within DiSIEM OTD [4], aiming to help analysts to prioritize the risk treatment and the response to potential incidents. Nonetheless, our proposal can target different OSINT sources and be applied to a large spectrum of TIPs. As described, the integration between different platforms and tools could be beneficial for the assessment and enrichment of threat intelligence, in an era where data is and comes from everywhere, and where having timely and accurate intelligence is determinant to protect an organization.

Our prototype uses as source a dataset collected on DiSIEM OTD [4] for 2 months, enriched with MCD [7], that is then parsed to allow the metrics computation. In total, we have assessed 487 tweets, submitted by 54 distinct accounts. The results from the experimental evaluation and from the conducted interviews show that our proposal can ease the risk treatment and favor the analysis of cyber threats.

We also identified improvements and features to include in future research works, that would extend DiSIEM OTD [4] capabilities and its visibility across the market.

7.1 Future work

During the implementation phase and experimental evaluation, we noticed several limitations on our work and some ideas have started to rise.

Implementing an API would allow external platforms to query accounts, tweets and IoCs scores. The *Twitter* API rate limit was not blocking during our experimental evaluation but it's certainly an nice-to-have in a tool like this, where huge amounts of data are processed. The lack of integration within DiSIEM OTD [4] is another point of improvement, that could allow to have the scores managed in a proper database solution. We got to know *CVE Trends* (Appendix C), a crowd-sourced CVE intelligence platform that uses an "audience" like metric (sum of all followers for each Twitter user that tweets or retweets a given CVE), that we consider that could to be a good addition to the proposed

set of metrics. The adoption of CPE (Common Platform Enumeration) would also be a huge improvement to this work, helping to know precisely which systems, software, and packages are impacted by a certain threat.

The interviewees also proposed some improvements and nice-to-have features, that we agree that would expand DiSIEM OTD [4] capabilities:

- Mapping with frameworks like *MITRE ATT&CK*;
- Give the ability to ignore certain *Twitter* accounts;
- Integration with CMDB to have a single asset management point and the ability to provide DiSIEM OTD [4] with more asset details;
- Keywords updating from highly reputable *Twitter* accounts;
- Refine D9 with asset exposition, by distinguishing between assets that are essential to the business and external assets.

Appendix A

Admiralty Grading System

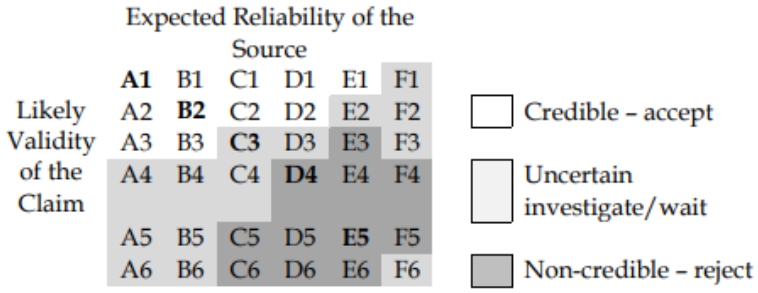


Figure A.1: NATO standard

	Reliability of the source		Credibility of the information
A	Completely reliable	1	Confirmed by other sources
B	Usually reliable	2	Probably true
C	Fairly reliable	3	Possibly true
D	Not usually reliable	4	Doubtful
E	Unreliable	5	Improbable
F	Reliability cannot be judged	6	Truth cannot be judged

Figure A.2: Reliability of the source (A-F) and credibility of the information (1-6)

Appendix B

NVD Vulnerability Severity Ratings

Severity	Base Score Range
Low	0.0-3.9
Medium	4.0-6.9
High	7.0-10.0

Figure B.1: CVSS v2.0 Ratings

Severity	Base Score Range
None	0.0
Low	0.1-3.9
Medium	4.0-6.9
High	7.0-8.9
Critical	9.0-10.0

Figure B.2: CVSS v3.0 Ratings

Appendix C

CVE Trends - crowdsourced CVE intelligence

Figure C.1: CVE Trends - web interface

Appendix D

Interviews - Questions

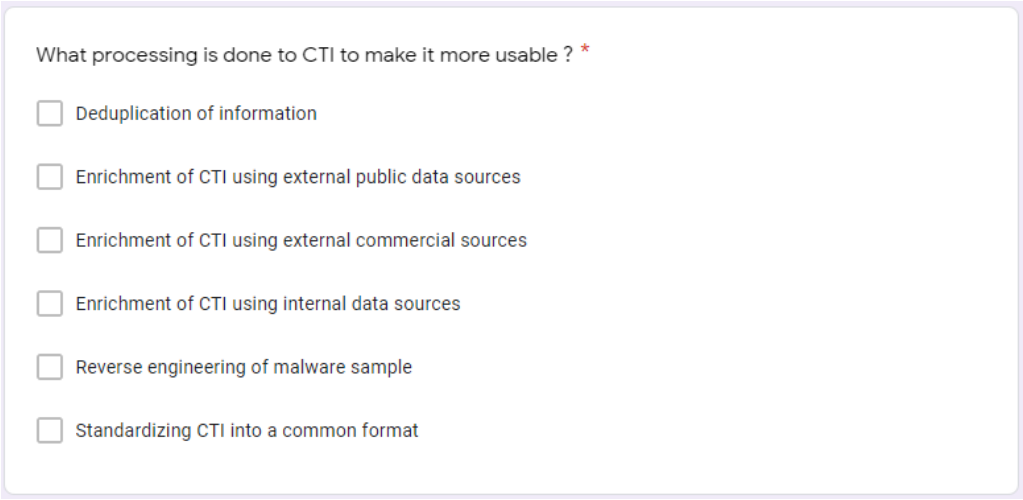


Figure D.1: Interviews - Question 1

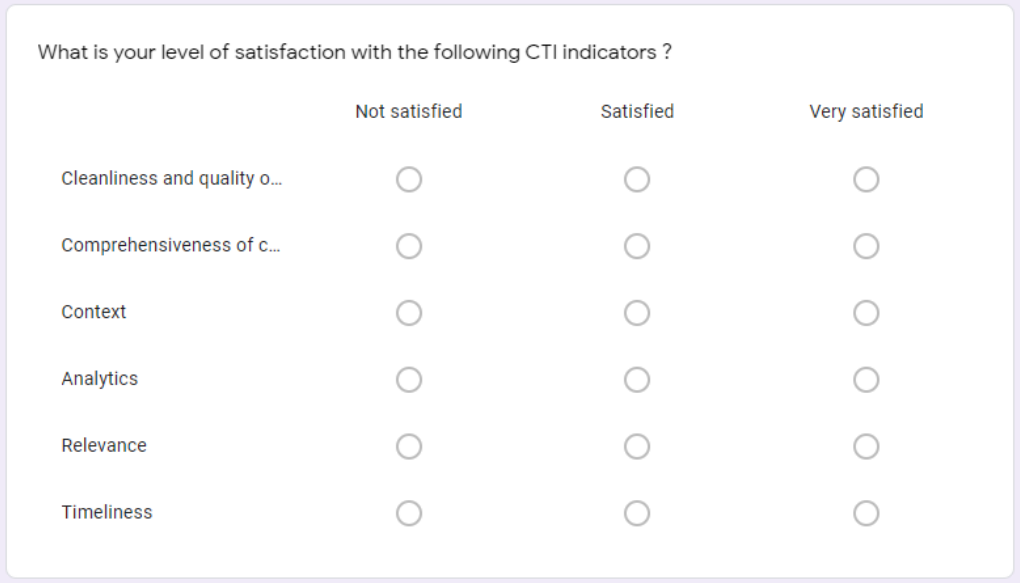


Figure D.2: Interviews - Question 2

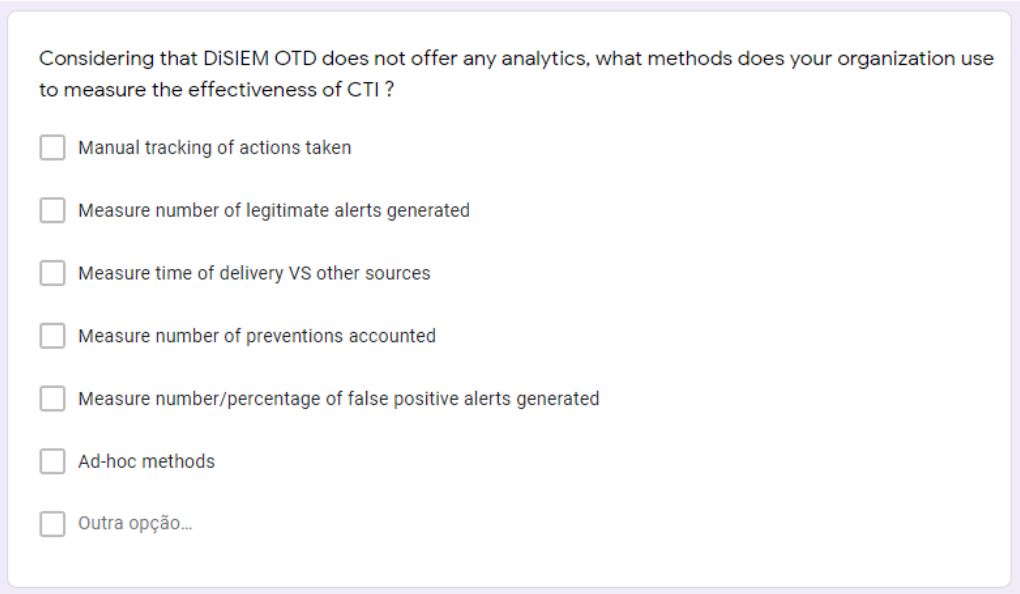


Figure D.3: Interviews - Question 3

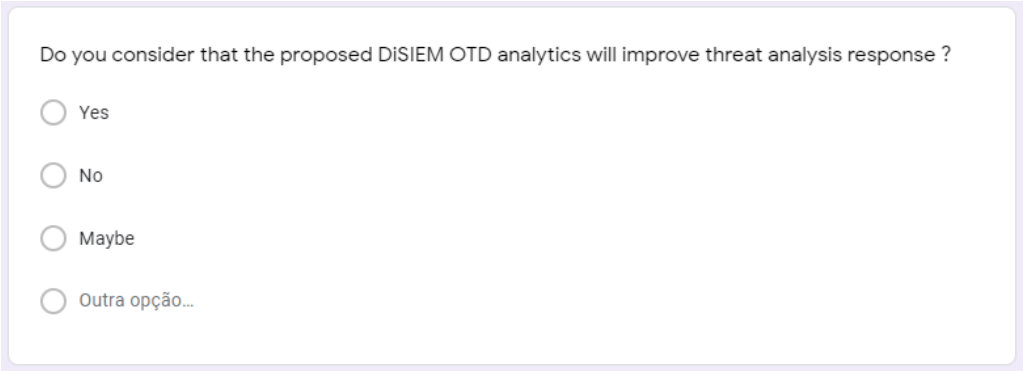


Figure D.4: Interviews - Question 4

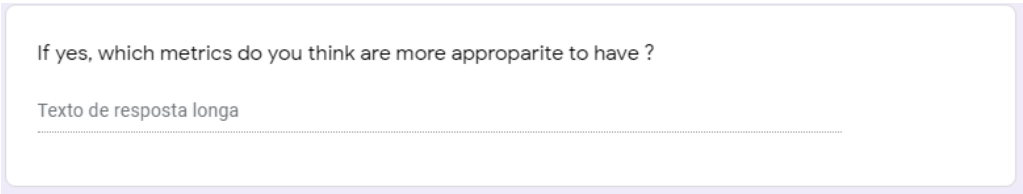


Figure D.5: Interviews - Question 5

Acronyms

API Application Programming Interface.

C2 / C&C Command and control.

CASE Cyber-investigation Analysis Standard Expression.

CIO Chief Information Officer.

CIRC Computer Incident Response Center.

CISO Chief Information Security Officer.

CMDB Configuration management database.

CSIRT Computer Security Incident Response Team.

CSV Comma-Separated Values.

CTI Cyber Threat Intelligence.

CTI Chief Technology Officer.

CVE Common Vulnerabilities and Exposures.

DNS Domain Name System.

DR Defense and response.

ENISA European Union Agency for Cybersecurity.

IAP In-App purchase.

ID Identifier.

IDS Intrusion Detection Systems.

IoC Indicator of Compromise.

IoT Internet of things.

IP Internet Protocol.

IPS Intrusion Prevention Systems.

IR Incident response.

JSON JavaScript Object Notation.

MCD Multitask Cyberthreat Detection.

MISP Malware Information Sharing Platform.

NATO North Atlantic Treaty Organization.

NCSC National Cyber Security Centre.

NOC Network Opertions Center.

NVD National Vulnerability Database.

OSINT Open Source Intelligence.

OTD OSINT Threat Detector.

RDF Resource Description Framework.

RSS Really Simple Syndication.

SCO STIX Cyber-observable Objects (SCO).

SIEM Security Information Event Management.

SOC Security Operations Center.

SRO STIX Relationship Object(SRO).

STIX Structured Threat Information Expression.

SVCE Security Vulnerability Concept Extractor.

TI Threat Intelligence.

TIP Threat Intelligence Platform.

TTL Time-to-live.

TTP Tactics, Techniques, Procedures.

UCO Unified Cyber Ontology.

URL Uniform Resource Locator.

USP User system profile.

VM Vulnerability managament.

XML eXtensible Markup Language.

Bibliography

[1] Cybersecurity Ventures, "Cybercrime To Cost The World USD 10.5 Trillion Annually By 2025." https://cybersecurityventures.com/cybercrime-damages-6-trillion-by-2021/.

[2] World Economic Forum, "Global Risks Report 2020 - Wild Wide Web." https://reports.weforum.org/global-risks-report-2020/wild-wide-web/.

[3] ENISA, "ENISA Threat Landscape 2020 - Cyber threat intelligence overview." https://www.enisa.europa.eu/publications/cyberthreat-intelligence-overview.

[4] LASIGE, "DiSIEM Project." https://disiem.lasige.di.fc.ul.pt.

[5] ISO, "ISO/IEC 25012:2008." https://www.iso.org/standard/35736.html.

[6] N. Dionísio, "Github code repository: multitask-cyberthreat-detection." ttps://github.com/ndionysus/multitask-cyberthreat-detection.

[7] N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, "Towards end-to-end cyberthreat detection from twitter using multi-task learning," in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, 2020.

[8] SANS, "2021 SANS Cyber Threat Intelligence (CTI) Survey." https://www.cybersixgill.com/wp-content/uploads/2021/02/SANS_CTI_Survey_2021_Sixgill.pdf.

[9] F. Alves, A. Bettini, P. M. Ferreira, and A. Bessani, "Processing tweets for cybersecurity threat awareness," 2020. Information Systems 2021.

[10] Cyber-Trust, "D2.2 Threat sharing methods: comparative analysis." https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bd41a4c2&appId=PPGMS.

[11] Kaspersky, "Evaluating threat intelligence sources." https://www.kaspersky.com/blog/evaluating-threat-intelligence/26952/.

[12] A. de Melo e Silva, J. J. Costa Gondim, R. de Oliveira Albuquerque, and L. J. García Villalba, "A methodology to evaluate standards and platforms within cyber threat intelligence," *Future Internet*, vol. 12, 2020.

[13] Kaspersky, "Advanced Protection and Threat Intelligence to Mitigate the Risk of Targeted Attacks." https://media.kaspersky.com/en/enterprise-security/threat-management-defense-solution-white-paper.pdf.

[14] D. F. SANS, "Threat intel processing at scale," 2019.

[15] CREST, "What is Cyber Threat Intelligence and how is it used?." https://www.crest-approved.org/wp-content/uploads/CREST-Cyber-Threat-Intelligence.pdf.

[16] H.-Y. L. SANS, "Using ioc (indicators of compromise) in malware forensics," 2013.

[17] David Bianco, "The Pyramid of Pain." https://rvasec.com/slides/2014/Bianco_Pyramid%20of%20Pain.pdf.

[18] CIRCL - Computer Incident Response Center Luxembourg, "MISP Default Feeds." https://www.misp-project.org/feeds/.

[19] OASIS CTI Technical Committee, "Introduction to STIX." https://oasis-open.github.io/cti-documentation/stix/intro.html.

[20] OASIS CTI Technical Committee, "STIX 2.1 Examples." https://oasis-open.github.io/cti-documentation/stix/examples.html.

[21] OpenIOC.org, "Sophisticated indicators forthe modern threat landscape: An introduction to openioc," 2012.

[22] market.us, "Twitter Statistics and Facts." https://market.us/statistics/social-media/twitter/.

[23] C. Sabottke, O. Suciu, and T. Dumitraş, "Vulnerability disclosure in the age of social media: Exploiting twitter for predicting real-world exploits," 2015.

[24] S. Mittal, P. Das, V. Mulwad, A. Joshi, and T. Finin, "CyberTwitter: Using Twitter to generate alerts for Cybersecurity Threats and Vulnerabilities," August 2016.

[25] W3, "Resource Description Framework." https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/.

[26] UCO Community, "Unified Cyber Ontology (UCO)." https://github.com/ucoProject/UCO.

[27] CASE Community, "Cyber-investigation Analysis Standard Expression (CASE)." https://caseontology.org/ontology/intro.html.

[28] A. Tundis, S. Ruppert, and M. Max, "On the Automated Assessment of Open-Source Cyber Threat Intelligence Sources," June 2020.

[29] D. Schlette, F. Böhm, M. Caselli, and G. Pernul, "Measuring and visualizing cyber threat intelligence quality," March 2020.

[30] R. Azevedo, I. Medeiros, and A. Bessani, "PURE: Generating Quality Threat Intelligence by Clustering and Correlating OSINT," August 2019.

[31] UK NCSC, "Cyber Threat Intelligence in Government: A Guide for Decision Makers & Analysts." https://hodigital.blog.gov.uk/wp-content/uploads/sites/161/2020/03/Cyber-Threat-Intelligence-A-Guide-For-Decision-Makers-and-Analysts-v2.0.pdf.

[32] CrowdStrike, "What is Threat Intelligence?." https://www.crowdstrike.com/cybersecurity-101/threat-intelligence/.

[33] TrendMicro, "Hunting threats on Twitter." https://www.trendmicro.com/vinfo/es/security/news/cybercrime-and-digital-threats/hunting-threats-on-twitter.

[34] ENISA, "ENISA Exploring the opportunities and limitations of current Threat Intelligence Platforms." https://www.enisa.europa.eu/publications/exploring-the-opportunities-and-limitations-of-current-threat-intelligence-platforms.

[35] C. Salvatore, S. Biffignandi, and A. Bianchi, "Social media and twitter data quality for new social indicators," Springer Netherlands, February 2020.

[36] L. Cai and Y. Zhu, "The challenges of data quality and data quality assessment in the big data era," *Data Science Journal*, vol. 14, 2015.

[37] F. Arolfo, K. C. Rodriguez, and A. Vaisman, "Analyzing the quality of twitter data streams," *Information Systems Frontiers*, 2020.

[38] T. Schaberreiter, V. Kupfersberger, K. Rantos, A. Spyros, A. Papanikolaou, C. Ilioudis, and G. Quirchmayr, "A quantitative evaluation of trust in the quality of cyber threat intelligence sources," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, ARES '19, (New York, NY, USA), Association for Computing Machinery, 2019.