IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 November 2023, accepted 4 January 2024, date of publication 8 January 2024,
date of current version 16 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3351108

RESEARCH ARTICLE

A Methodology for Developing & Assessing CTI Quality Metrics

GEORGIOS SAKELLARIOU, PANAGIOTIS FOULIRAS, AND IOANNIS MAVRIDIS
Department of Applied Informatics, School of Information Sciences, University of Macedonia, 546 36 Thessaloniki, Greece

Corresponding author: Panagiotis Fouliras (pfoul@uom.edu.gr)

The publication of the article in OA mode was financially supported by HEAL-Link.

ABSTRACT Since its first steps in the cybersecurity field, Cyber Threat Intelligence (CTI) has gained recognition and increased its importance in the daily operations of cybersecurity teams. However, the many forms of CTI exchanged, the vast amount of CTI products, and the plurality of the sources have raised doubts about the CTI quality. This paper discusses the problem of CTI quality, focusing on the quality factors that better evaluate the products of CTI and how we measure them. Consequently, we propose a methodology for developing and assessing CTI quality metrics and demonstrate the application of this methodology by developing the relevance (RE) and weighted completeness (WC) metrics for unstructured and structured CTI products, respectively. We created two sets of structured and unstructured CTI data for this demonstration, utilizing them as benchmark datasets for estimating RE and WC.The proposed methodology introduces a systematic approach for developing and assessing quantitative CTI quality metrics for evaluating CTI data and CTI sources.

INDEX TERMS Cyber threat intelligence, cybersecurity, information sharing, datasets, quality, similarity.

I. INTRODUCTION

Cyber Threat Intelligence (CTI) is a new field of cybersecurity with almost a decade of presence. However, an increasing number of organizations have adopted the benefits of CTI to safeguard their security posture. The CTI process exploits information gathered from various sources to provide ''intelligence'' for an organization. Then, the organization's security experts use this ''intelligence'' to detect, prevent, or predict a cyber attack and for their decision-making [1]. For consistency, in this paper, we call the various forms of ''intelligence'', *CTI products*, when produced by a CTI process, either within the organization or from an external entity. We call all gathered information from *CTI sources*, *''CTI data''*, which may be CTI products produced by external entities, *raw CTI data* (non-CTI products), or a combination of them [1].

In general, security experts have to deal with a vast amount of CTI data and CTI sources that may have minor or no meaning to them [2]. Additionally, other related research, for example, on the trust of CTI sources and the actionability of CTI products, are also mentioned in the literature [3], [4]. Those issues related to CTI sources, CTI data, or the processes applied to them can be seen under the umbrella of *CTI quality* [5]. CTI quality and its factors have been discussed in several papers (e.g., [4]). Although determining and following a qualitative approach to evaluating these factors does exist in the bibliography [6], a lack of relevant research appears in their quantitative measurement.

This paper aims to address the quantitative measurement of CTI quality by researching the ''CTI'' based on the following questions:

1) Q1: Which quality factors affect the evaluation of CTI data and CTI sources more?
2) Q2: How can we define metrics for the CTI quality factors?
3) Q3: Can we use these metrics to assess the quality of large amounts of CTI data in a timely manner?

This study answers these questions by (a) discussing and analyzing the background behind the CTI quality factors and

The associate editor coordinating the review of this manuscript and approving it for publication was Mansoor Ahmed.

the characteristics of CTI data and CTI sources, (b) proposing a methodology for CTI quality metrics development and assessment, and applying it in defining two CTI quality metrics, and (c) applying those metrics on datasets of structured and unstructured CTI products and comparing with the results of other proposed CTI quality metrics in the bibliography.

The remainder of this paper is organized as follows. Section II analyzes the existing work on CTI quality by conducting bibliography research. Section III discusses and analyzes the background of CTI quality, focusing on the relation between CTI quality factors and the characteristics of CTI data and CTI sources. It also identifies the CTI quality factors that primarily affect the evaluation of CTI data and CTI sources. Section IV presents our CTI quality metrics development and assessment methodology, and Section V defines two metrics by applying the proposed methodology. The use of those metrics and their comparison with the results of other metrics proposed in the literature is the subject of Section VI. Section VII, discusses the relation of the proposed methodology with other aspects of CTI and future research initiatives. Finally, Section VIII summarizes the contributions of this study.

II. RELATED WORK

Although CTI is a new cybersecurity area, some scientific publications have attempted to identify the factors related to CTI quality and define methods and metrics to measure them. Next, we present an overview of these studies.

Schlette et al. [7] considered the objective and subjective CTI quality dimensions, further dividing them into the attribute, object, and report levels. They then identified and defined the estimation of nine quality metrics based on the characteristics of STIX v2. However, the authors did not analyze the quality aspects of the CTI sources, focusing only on measuring the quality of CTI products.

Qiang et al. [8] proposed an evaluation architecture to measure the quality of what they call CTI services (including support services) offered by various vendors from a user's perspective. They present an indexing system that examines a set of quality characteristics (e.g., price of data and reputation) and employs quantization techniques, normalization methods, and attribute weights. However, the authors did not discuss how to assess the quality of CTI products (e.g., timeliness and accuracy).

Schaberreiter et al. [9] proposed a methodology to estimate the quality of a CTI source by measuring parameters (quality factors) such as the similarity, completeness, and timeliness of the CTI products. Their overall scope is to create trust indicators for a CTI source's quality, which continuously reflects a CTI source's trust level compared to others. However, this methodology focuses only on evaluating the quality of the CTI source and does not provide an independent estimation of the CTI product's quality.

Furthermore, Meier et al. [10] proposed an algorithm that ranks cyber threat intelligence feeds (essentially CTI products) quality using a similar approach to PageRank but applied to a CTI feeds graph and takes as input: (a) the completeness (measures the contribution of a CTI feed to the set of all entries), (b) the accuracy (measure the entries confirmed by other feeds), and (c) the speed (the number of the entries entered by a feed on a time unit) quality factors. This algorithm aims to create a quality-based ranking of the CTI sources; however, the authors do not discuss how to measure the quality of individual CTI products.

Zibak et al. [6] determined the essential quality factors of CTI products and CTI sources (accuracy, actionability, interoperability, provenance, relevance, reliability, and timeliness) after performing an extensive literature review and a Delphi-based [11] study. However, they did not discuss how to measure those quality factors. On the other hand, in [12], Martins and Medeiros neither tried to identify the quality factors of CTI nor measure them; instead, they proposed a method that improves the quality by automatically classifying threat intelligence data and enriching them with OSINT.

During the FIRST 2015 symposium [13], Pinto and Sieira, proposed using a data-driven set of tests (TIQ, novelty, aging, population, overlap, and uniqueness tests) to estimate the quality of CTI products without investigating the quality characteristics demanded by the security analysts.

Additionally, in the FIRST 2016 symposium [14], a simplified methodology was proposed to measure the quality of CTI feeds by estimating the quality factors of delivery delay (which supposedly reflects their timeliness), false-positive rate, cross-dataset linkage (existence of the same CTI product on different sources), and utility (number of researchers' queries on each feed). Although this methodology comes from the author's practical experience, it has not yet been evaluated.

Finally, in [15], Zhang et al. evaluated the quality by counting the number of Indicators of Compromise (IoCs) extracted from the CTI products and the number of attack techniques revealed because of the discovered IoCs instead of measuring the quality of CTI products, quantitatively.

By examining the literature on CTI quality, we identified the gap between the theoretical definition of quality factors and the actual estimation of the quality metrics that reflect them. This study examines the identified gap and proposes a methodology that can produce and assess metrics for any relevant CTI quality factor or a combination of them.

III. BACKGROUND

A. DEFINITIONS

By studying the CTI quality-related bibliography, we also tracked confusion on how the terms factors, metrics, and measurement methods are used. Therefore, to use these terms unambiguously in this study, we use the term *quality factor(s)* to refer to abstract quality concepts such as timeliness and accuracy. We also use the term *quality metrics* when we refer to quantifying those quality factors or any combination of them in the context of specific CTI products or sources. Finally, we use the term *measurement methods* when referring to the calculation methodologies applied to quality metrics.

B. QUALITY FACTORS

To measure CTI quality, we need first to identify those quality factors that affect CTI and second to clarify the entities of CTI to which quality factors are referred.

In the literature, we identified three areas of CTI that have a specific relation with at least one quality factor. We used these areas to categorize and interrelate the quality factors with the identified CTI entities. In Table 1, we enumerate the quality factors related to CTI and categorize them into three areas: the area of *collected data*, the area of *produced intelligence*, and the area of *information sharing*. We observe that a quality factor may belong to more than one area because the same term is used to express slightly different concepts (e.g., the accuracy of metrics and accuracy).

TABLE 1. CTI quality factors.

CTI area of	Quality Factors
collected data	accuracy, timeliness, completeness, consistency [16], relevance, actionability, value [17].
produced intelligence	accuracy [3], [4], [18]–[20], clarity, utility [18], relevance [3], [4], [20], [21], timeliness [3], [4], [19]–[21], actionability [3], [4], [19], completeness [3], [19], [20], ingestibility, trustworthiness [3]
information sharing	consumer-based evaluation [4], total quality of shared information [22], traceability, provenance [21], uncertainty of sharing [23]

Based on the literature review outlined in Section II, we realized a tendency to express CTI quality exclusively as the quality of CTI sources. However, this leads to ignoring other important entities in which CTI quality should be measured. Generally, as mentioned in Section I, we can distinguish the following broad categories of entities: CTI data (raw CTI data, CTI products) and CTI sources.

Raw CTI data are unrefined data used as input by a CTI process and have various forms and features (e.g., logs, security events, discussion forums, malware analysis data, geo-location data, etc.). Raw CTI data are associated with the CTI area of collected data. Given that raw CTI data are the input of a threat intelligence process, we should consider that their quality factors, as presented in the area of collected data in Table 1, are constant and unaffected by the CTI process.

CTI products are threat intelligence process outputs that can take a structured (e.g., STIX, OpenIOC, YARA) or unstructured (e.g., security report) format. Therefore, the quality of CTI products is generally affected by the factors listed in Table 1 in the CTI area of produced intelligence.

As CTI products are the outcome of a systematic process, the CTI quality factors listed in Table 1 should be met. However, the unstructured and structured forms of CTI products significantly impact how we can define a metric of these quality factors. Therefore, we can distinguish the metrics of CTI products into those calculated for unstructured CTI products and structured CTI products.

Furthermore, measuring the quality of raw CTI data or CTI products does not necessarily mean we can infer the quality of CTI sources. The assumption that the accumulated quality of CTI products reflects the overall quality of the source seems obvious, especially if we consider a CTI source as a collection of CTI products, but it is not the only one. Instead, CTI information sharing is the fundamental difference between a CTI product's collection and a CTI source [3]. Accordingly, we present the CTI information-sharing quality factors that can reflect the overall quality of a CTI source in Table 1 (e.g., traceability).

Till this point, we have presented the three areas of CTI quality factors of Table 1. The quality metrics aim to quantify these CTI quality factors; however, in the literatures [1] and [24], we observe that quality factors also exist for quality metrics. These quality factors are objectivity, subjectivity, performance, behavior, and accuracy of metrics. We later handle the quality factors referring to the quality of metrics as part of the proposed methodology.

IV. METHODOLOGY OF CTI QUALITY METRICS DEVELOPMENT AND ASSESSMENT

A metric can be considered a tool for decision-making that quantitatively measures a specific notion [24], in our case, the CTI quality factors or any combination of these factors. However, as explained in Section III-B, even metrics are affected by quality factors. Thinking of a metric as a tool makes it easier to explain the role and the impact of the quality factors in the creation of a CTI quality metric (M): objectivity (O), subjectivity (S), performance (P), behavior (B), and accuracy (A). These quality factors of metrics correspond to the four categories of cyber security measurements of situational awareness [25] plus the accuracy, as described in [24]. We introduce these quality factors in the development of CTI quality metrics by leveraging the strong relation between cybersecurity situational awareness (SA) and CTI [26], [27]; besides, they express measurement quality independently from cybersecurity [28].

By examining the above factors, we observed that subjectivity and objectivity are inversely proportional concepts that indicate human involvement in a process [29], specifically a measurement process in our case. Of course, measuring the subjectivity and objectivity of a process is complicated and can even touch on philosophical aspects [30]. In the case of CTI, we need to indicate a metric's level of subjectivity and objectivity effectively. A lack of related work in cyber security leads us to examine the topic from the point of view of other scientific areas. In [31], Rothstein described an objective measurement process as one that always gives the same result when applied to the same data, independent of the person who uses it. He also distinguished between the subjectivity and objectivity of a metric and the subjectivity and objectivity of the data itself (e.g., a metric that counts the number of positive votes of cyber security experts about the value of a CTI source is objective because it only counts the votes, but the votes themselves are subjective because they involve the human factor). Based on this observation, the author distinguished four subjectivity/objectivity levels for a metric: (a) a subjective metric of objective data (SO), (b) a subjective metric of subjective data (SS), (c) an objective metric of subjective data (OS), and (d) an objective metric of objective data (OO). In this work, we adopt these levels to determine a CTI metric's objectivity and subjectivity, represented as (Γ), and present how we use them in the proposed methodology.

The measurement of a metric is typically performed by executing an algorithm. Therefore, we consider a metric's *performance* (P) quality factor to coincide with the performance of this algorithm, particularly with its time complexity.

In the SA, behavioral metrics infer the level at which the SA is affected by individual actions [26], [27]. From a CTI point of view, this seems generic; instead, a factor that reflects the strength of a CTI quality metric against adversaries' behavior is necessary because CTI data are often the target of malicious activities [32]. We adopt the "behavior" quality factor, which reflects the strength of a metric against the attacks that modify the CTI data (i.e., data from CTI entities) used by the metric. We define the behavior quality factor as describing a metric's sensitivity against an adversary's actions on the CTI entities. However, since characterizing someone as an adversary may not be feasible at this point, we can say that the behavior quality factor is equal to the sensitivity of a metric to a change without determining whether the change is adversarial. Furthermore, to determine the behavior quality factor, we perform a sensitivity analysis [33] in which we evaluate how a given metric M, calculated by employing a set of variables X, is affected by changes in X. The bibliography includes various sensitivity analysis methods [34]. However, in this study, we avoid proposing any of them as the most appropriate for estimating the behavior quality factor of CTI metrics, considering it as a future research challenge. We assume that the behavior quality factor is represented by variable B, which is the set of sensitivity analysis results.

The last quality factor, accuracy, appears only in the non-CTI bibliography. Accuracy determines the level at which a metric's estimation agrees with the actual metric's value. The difference between this estimation and the respective real value is called *bias* (b). In the case of CTI quality metrics, we assume function $A(b)$, expressing a metric's accuracy as a value between 0 and 1. In our case, bias represents the uncertainty introduced by a metric in the measurement process (e.g., by rounding a value). Let M be a CTI quality metric calculated on a set of variables X, $M = F(X)$, where M_r is the actual value of M and M_c is the calculated value of M, and let b be the bias introduced by the calculation of this metric. Then, the real value of M (M_r) is given by the equation $M_r = A(b)M_c$. In the simplest case, the accuracy $A(b)$ is expected to be one. However, the accuracy will be more pronounced in the future because we expect to define complex metrics by following a probabilistic approach in CTI. For example, a metric expresses the probability that a CTI product is relevant to an attack vector in the short term.

Generally, we define a CTI quality metric as a tuple of both the set of the previous factors $Q = \{\Gamma, P, B, A\}$ and a function F that acts on a set of variables X.

$$M = (Q, F(X)) \tag{1}$$

where:

$$Q = \{\Gamma, P, B, A\},\ \ \Gamma \,\epsilon \{SO, SS, OS, OO\},$$
$$P = \{expressed\ as\ time\ complexity\},$$
$$B = \{results\ of\ sensitivity\ analysis\},\ \ A = A(b)\,\epsilon [0, 1]$$

Thus, we propose a CTI quality metrics development and assessment methodology that not only constructs metrics that measure a CTI entity's quality factors but also offers a way to assess and compare these metrics.

The proposed methodology comprises eight steps and has zero "quality-related" knowledge of CTI as a prerequisite. Next, we briefly describe these steps.

- Step 1. We examine the raw CTI data, or CTI products, or sources for which we want to estimate their quality, considering the quality factors of Table 1 and aiming to identify what better expresses their quality. More specifically, Table 1 can guide the development of a CTI quality metric by helping a researcher answer the following questions: (a) For what CTI area do we want to develop a CTI quality metric? (b) What CTI quality factors have already been proposed in the literature for this CTI area? For example, let us assume that we want to measure how up-to-date a CTI product (produced intelligence) is. Table 1 allows us to identify 'timeliness' as the appropriate CTI quality factor. After having identified these quality factors, we name a metric that can express them (e.g., for timeliness) metric M.
- Step 2. We determine the set of variables X, with which we calculate M (e.g., in the case of the timeliness metric, it could be a set of different timestamps of a CTI product).
- Step 3. We define the function F, which computes the M (e.g., in the case of the timeliness metric, it can be the average of the CTI product's timestamps).
- Step 4. We analyze X and F to determine subjectivity and objectivity Γ. The involvement of a human factor indicates subjectivity, while determinism indicates objectivity. Consequently, two questions guide the development of a CTI metric in this step: (a) Is a human factor involved in determining X? (b) is F deterministic?

- Step 5. We analyze F to determine M's performance P. For example, we perform a time complexity analysis of the algorithm that calculates F.
- Step 6. We determine M's accuracy A based on whether the algorithm employed for calculating F is deterministic or not. If F is calculated by a deterministic algorithm without rounding, then $A = 1$. In case of a non-deterministic algorithm, A is calculated experimentally, as in a machine learning assisted approach.
- Step 7. We conduct a sensitivity analysis by applying one of the available methods in the literature [33] to determine the behavior B of the M.
- Step 8. We construct the tuple of metric $M = (Q, F(X))$.

V. APPLICATION OF THE PROPOSED METHODOLOGY

To apply the proposed methodology to the development of CTI quality metrics, let's assume that organization C operates on a finite set of industrial domains $D = \{d_1, d_2, \ldots., d_l\}$, with $l \in \mathbb{N}$. Organization C has a finite set of information technology assets $I = \{i_1, i_2, \ldots., i_g\}$, with $g \in \mathbb{N}$. Additionally, we assume source S_u of unstructured CTI products $P_u = \{p_{u1}, p_{u2}, \ldots, p_{un_1}, \}$ (e.g., CTI reports), and source S_s of structured CTI products $P_s = \{p_{s1}, p_{s2}, \ldots, p_{sn_2}, \}$ (e.g., CTI STIX v2.1 artifacts) with $n_1, n_2 \in \mathbb{N}$.

A. RELEVANCE OF UNSTRUCTURED CTI PRODUCTS

Following the proposed methodology (see Section IV), we develop a metric that measures the *relevance* (cf. Table 1) quality factor of unstructured CTI products. This quality factor expresses the level at which a CTI product is related to an entity, in our case, an organization.

1) STEP 1

Analyzing the essential facts for this metric, we have organization C and source S_u of unstructured CTI products (P_{un}). In addition, we assume the quality metric RE (metric under development) that measures the relevance of P_{un} concerning C.

2) STEP 2

To determine the set of variables X used for calculating RE, we observe that, on the one hand, we have C's information technology assets I and the industrial domains D. On the other hand, P_{un} has an unstructured text format. So, we conclude $X = \{I, D, P_{un}\}$. Before proceeding to the next step, we need to clarify what those variables represent in the real world or what security experts use to represent the notions behind those variables. An organization's information technology assets I can be described and enumerated in various ways. A commonly accepted approach is the use of the *Common Platform Enumeration* (CPE) scheme [35], with a total of eleven keys by default. For our purpose, we match an asset (i_g) of organization C with a CPE entry. A CPE entry comprises a key-value pairs dictionary $\{k_1 : v_1, \ldots, k_{11} : v_{11}\}$, as presented in Table 2. Next, to define D, we use the Domain Industry Taxonomy (DIT) [36]; specifically, we consider that C's industrial domain (d_l) is described by a DIT *main category* (e.g., manufacturing), and for each d_l, we construct a set of words $\{dw_{l1}, dw_{l2}, \ldots, dw_{lz}\}$, $z \in \mathbb{N}$ (e.g., {pulp, paper, paperboard}) based on the DIT NACE *category names*. Hence, we have $D = \{d_1, d_2, \ldots., d_l\} = \{\{dw_{11}, dw_{12}, \ldots, dw_{1a}\}, \{dw_{21}, dw_{21}, \ldots, dw_{2m}\}, \ldots, \{dw_{l1}, dw_{l2}, \ldots, dw_{lz}\}\}$, where $a, m, z \in \mathbb{N}$. Finally, P_{un} is a text document consisting of a set of words (terms), w_i, and their respective frequency of appearance, f_i. Hence, $P_{un} = \{(w_1, f_1), (w_2, f_2), \ldots, (w_i, f_i)\}, i \in \mathbb{N}$. In summary, we have determined the set of variables $X = \{I, D, P_{un}\}$ as follows:

$$\begin{cases} I = \{i_1, i_2, \ldots., i_g\} = \{\{k_1 : v_1, \ldots, k_{11} : v_{11}\}_1, \ldots, \\ \{k_1 : v_1, \ldots, k_{11} : v_{11}\}_g\}, g \in \mathbb{N} \\ D = \{d_1, d_2, \ldots., d_l\} = \{\{dw_{11}, dw_{12}, \ldots, dw_{1a}\}, \\ \{dw_{21}, dw_{21}, \ldots, dw_{2m}\}, \ldots, \\ \{dw_{l1}, dw_{l2}, \ldots, dw_{lz}\}\}, a, m, z \in \mathbb{N} \\ P_{un} = \{(w_1, f_1), (w_2, f_2), \ldots, (w_i, f_i)\}, i \in \mathbb{N} \end{cases} \tag{2}$$

TABLE 2. CPE entry description.

Key	Key Name	Value Description
1	*part*	*a*-applications, *o*-operating systems, *h*-hardware devices
2	*vendor*	the manufacturer
3	*product*	title or name of the product
4	*version*	version of the product
5	*update*	update, service pack, or product release
6	*edition*	ANY
7	*language*	language tags described in RFC5646
8	*sw_edition*	characterize the product to a particular market
9	*target_sw*	software environment in which the product operates
10	*target_hw*	instruction set architecture
11	*other*	general information

3) STEP 3

To define function F, we observe that I, D, P_{un} can be representations of text documents. Therefore, we can transform the CTI quality metric estimation problem into a documents similarity estimation problem with $F_1 = \mathit{similarity}(I, P_{un})$ and $F_2 = \mathit{similarity}(D, P_{un})$. Because F1 and F2 are independent, we have $F(X) = F_1 \cdot F_2$. In the case of F_1, we use the cosine similarity [37], hence $F_1(I, P_{un}) = cos(d_I, d_{P_{un}}) = \frac{d_I \cdot d_{P_{un}}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}}\right\|}$, where d_I and $d_{P_{un}}$ are the document vector transformations of I and P_{un}, respectively. To perform the transformation $I \rightarrow d_I$, we count the frequencies of different values of each key across the CPE entries (e.g., k_1 =(part) → {(applications:3, hardware devices:5, operating systems:4}); and then we perform the union of those *value:frequency* pairs, thus forming a joint set $I' = \cup(v_j, f'_j), j \in \mathbb{N}$. Next, we construct a j-dimensional vector $d_I = \langle f'_1, \ldots., f'_j \rangle$. To construct $d_{P_{un}}$, we assume the temporary j-length zero vector $temp = \langle 0, 0, \ldots, 0 \rangle$, and for each word $w_i \epsilon P_{un}$, if w_i matches v_j, we set $temp(j) = f_i$; otherwise, we discard w_i. Finally, we set $d_{P_{un}} = temp$.

In the case of F_2, we use the Jaccard similarity [37], $F_2(D, P_{un}) = Jaccard(d_D, d'_{P_{un}}) = \frac{d_D \cap d'_{P_{un}}}{d_D \cup d'_{P_{un}}}$ where d_D and $d'_{P_{un}}$ ($d'_{P_{un}} \neq d_{P_{un}}$) are the document vector transformations of D and P_{un}, respectively, which is most appropriate for the comparison of asymmetric vectors as we expect to be the d_D and $d'_{P_{un}}$; wherein, we need to reveal at what level P_{un} is related to a set of industrial domains, represented by D. To construct d_D, we perform the transformation $D \rightarrow d_D : \bigcup \bigcup_{i=1}^{l} d_i$, which means that d_D is the set of the unique words, dw_{lz}, of D, with $min(a, m, .., z) \leq |d_D| \leq \sum(a, m, .., z)$. To construct the $d'_{P_{un}}$, we assume the temporary $|d_D|$-length zero vector $temp = \langle 0, 0, \ldots, 0 \rangle$. For each of the words $w_i \epsilon P_{un}$, if w_i matches in d_D, we add it to the first available zero position of $temp$. Finally, we set $d'_{P_{un}} = temp$. We observe that $d'_{P_{un}} \subseteq d_D$, so F_2 takes the form $F_2(D, P_{un}) = \frac{d_D \cap d'_{P_{un}}}{d_D \cup d'_{P_{un}}} = \frac{\left|d'_{P_{un}}\right|}{\left|d_D\right|}$. In summary, to avoid losing information about the relevance of P_{un}, if one of the functions is equal to zero, we have:

$$F = \begin{cases} F_2 \text{ , } if \quad F_1 = 0 \ \& \ F_2 \neq 0 \\ F_1 \text{ , } if \quad F_2 = 0 \ \& \ F_1 \neq 0 \\ F_1 \cdot F_2 \end{cases} \tag{3}$$

$$where\ F_1(I, P_{un}) = \frac{d_I \cdot d_{P_{un}}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}}\right\|}\ and\ F_2(D, P_{un}) = \frac{\left|d'_{P_{un}}\right|}{\left|d_D\right|}$$

4) STEP 4

To determine *subjectivity* and *objectivity*, Γ, we observe that one of the variables, D, is produced by the involvement of a human factor, which decides the industrial domains of organization C. Therefore, part of the data in which F is computed is considered subjective. Moreover, F results from the multiplication of F_1 and F_2, which are applications of cosine and Jaccard similarity functions, respectively. Because cosine and Jaccard functions are deterministic, we can deduce that F is objective. In summary, we infer that $\Gamma = OS$.

5) STEP 5

To estimate the performance of *RE*, using the proposed Algorithm 1 (see below), which calculates F, we observe that F is a product of F_1 and F_2 in the worst case. Additionally, F_1 and F_2 are independent and can be calculated in parallel. Hence, in the worst case, the *performance P* is the worst performance of those of F_1 and F_2. In the case of F_1, we have an application of cosine similarity, with $O(n)$ performance [38]. In the case of F_2, we have Jaccard similarity with $O(n^2)$ performance. However, because we avoid the logical union and intersection operations, F_2 results in $O(1)$ performance. In conclusion, we have $P = O(n)$ for the *RE* metric.

Algorithm 1 *RE* Metric Algorithm

```
Require: $d_I = [f'_1, \ldots, f'_j]$,
    $d_{P_{un}} = [f_1, \ldots, f_j]$, $d_D = [dw_{11}, \ldots, dw_{lz}]$, $d'_{P_{un}} = [w_1, \ldots, w_i]$
    $F, F_1, F_2, a, b, sum_1, sum_2 \leftarrow 0$
    for $k$ in $d_I$ do
        $a+ = k^2$
    end for
    for $k$ in $d_{P_{un}}$ do
        $b+ = k^2$
    end for
    $sum_1 = \sqrt{a} \cdot \sqrt{b}$
    for $i$ inrange $length(d_{P_{un}})$ do
        $sum_2+ = d_I[i] \cdot d_{P_{un}}[i]$
    end for
    $F_1 \leftarrow sum_2/sum_1$
    $F_2 \leftarrow count(d'_{P_{un}})/count(d_D)$
    if $F_1 = 0 \ \& \ F_2 \neq 0$ then
        $F \leftarrow F_2$
    else if $F_1 \neq 0 \ \& \ F_2 = 0$ then
        $F \leftarrow F_1$
    else
        $F \leftarrow F_1 \cdot F_2$
    end if
```

6) STEP 6

The calculation of *RE* follows a deterministic algorithm with no rounding or probabilistic procedures (see Algorithm 1). As explained in Section IV, F does not introduce *bias* in the calculation of *RE*; thus, $A = 1$.

7) STEP 7

To perform a sensitivity analysis of F, we first need to clarify those variables of X in which a potential change is not controlled by organization C. This distinction is made because in CTI, as we have already mentioned, we care about the behavior of a CTI metric on changes caused by potential adversaries. In our case, we observe that the only variable of X that can be affected by others except organization C is P_{un}. P_{un} participates with $d_{P_{un}} = \langle f_1, \ldots, f_j \rangle$ and vector size of $\left|d'_{P_{un}}\right|$ in the calculation of F. We consider vectors $\overline{\Lambda} = \langle \lambda_1, \ldots, \lambda_j, \rangle \ni 0 \leq \lambda_j \leq 1$ and $\overline{\Theta} \ni 0 \leq \left|\overline{\Theta}\right| \leq \left|d_D\right|$ the results of the change caused on P_{un} and affecting $d_{P_{un}}$ and $\left|d'_{P_{un}}\right|$, respectively. Based on that, we perform a sensitivity analysis on F following the *elementary effects method* [34].

To apply a sensitivity analysis using the *elementary effects method* [34] on *RE* as calculated by F, we observe that F is the conditional result of the multiplication of $F_1(I, P_{un}) = F_1(d_I, d_{P_{un}}) = \frac{d_I \cdot d_{P_{un}}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}}\right\|}$ and $F_2(D, P_{un}) = F_2(d_D, d'_{P_{un}}) = \frac{\left|d'_{P_{un}}\right|}{\left|d_D\right|}$, $F = F_1 \cdot F_2$. Therefore, because F_1 and F_2 are independent, we need to apply the method twice, considering once F_2 and then F_1 constant.

Case 1: Sensitivity analysis of F_1 keeping F_2 constant. Following the *elementary effects method* [34] for a group of factors, we consider p_α selected levels in which $d_{P_{un}}$ and $\overline{\Lambda}$ can be set, where the Ω_α is the respective discretized input space (i.e., the discrete vectors that $\overline{\Lambda}$ can be). Then, the elementary effect of $d_{P_{un}}$ in F_1 is:

$$EE_{d_{P_{un}}} = \frac{F_1(d_I, \overline{\Lambda}) - F_1(d_I, d_{P_{un}})}{\Delta_\alpha} \tag{4}$$

where $\Delta_\alpha \in \{\frac{1}{p_\alpha - 1}, 1 - \frac{1}{p_\alpha - 1}\}$ and $\lambda_j = f_j \pm \Delta_\alpha$. Then the distribution $F_{d_{P_{un}}}$ of $EE_{d_{P_{un}}}$ is derived by randomly sampling $\overline{\Lambda}$ from Ω_α.

According to theory [34], the sensitivity measures are the mean (μ) and the standard deviation (σ) of distribution $F_{d_{P_{un}}}$, and the mean of the absolute values (μ^*) of $\left|EE_{d_{P_{un}}}\right|$ of the respective distribution $\left|EE_{d_{P_{un}}}\right| \sim G_{d_{P_{un}}}$. Here, μ assesses the influence of $d_{P_{un}}$ in F_1 (i.e., in RE since F_2 is constant), μ^* assesses again the influence of $d_{P_{un}}$ in F_1 (i.e., RE) while simultaneously handling negative valued $EE_{d_{P_{un}}}$, and σ reveals the total effects of the interactions between $d_{P_{un}}$ and d_D.

Applying the theory's sampling strategy, we conclude that for the distributions $F_{d_{P_{un}}}$, $G_{d_{P_{un}}}$ derived from r_α samples, we have:

$$\begin{aligned}
\mu_{d_{P_{un}}} &= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} EE_{d_{P_{un}}i} = \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \frac{F_1(d_I, \overline{\Lambda_i}) - F_1(d_I, d_{P_{un}i})}{\Delta_{\alpha i}} \\
&= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \frac{\frac{d_I \cdot \overline{\Lambda_i}}{\left\|d_I\right\| \cdot \left\|\overline{\Lambda_i}\right\|} - \frac{d_I \cdot d_{P_{un}i}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i}\right\|}}{\Delta_{\alpha i}} \\
&= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \frac{\frac{d_I \cdot (d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i})}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i}\right\|} - \frac{d_I \cdot d_{P_{un}i}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i}\right\|}}{\Delta_{\alpha i}}
\end{aligned} \tag{5}$$

$$\begin{aligned}
\mu^*_{d_{P_{un}}} &= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \left|EE_{d_{P_{un}}i}\right| \\
&= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \frac{\left|F_1(d_I, \overline{\Lambda_i}) - F_1(d_I, d_{P_{un}i})\right|}{\Delta_{\alpha i}} \\
&= \frac{1}{r_\alpha}\sum_{i=1}^{r_\alpha} \frac{\left|\frac{d_I \cdot (d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i})}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i}\right\|} - \frac{d_I \cdot d_{P_{un}i}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i}\right\|}\right|}{\Delta_{\alpha i}}
\end{aligned} \tag{6}$$

$$\begin{aligned}
\sigma^2_{d_{P_{un}}} &= \frac{1}{r_\alpha - 1}\sum_{i=1}^{r_\alpha} \left(EE_{d_{P_{un}}i} - \mu_{d_{P_{un}}}\right)^2 \\
&= \frac{1}{r_\alpha - 1}\sum_{i=1}^{r_\alpha} \left(\frac{\frac{d_I \cdot (d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i})}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i} \pm [\mathbf{1}]\Delta_{\alpha i}\right\|} - \frac{d_I \cdot d_{P_{un}i}}{\left\|d_I\right\| \cdot \left\|d_{P_{un}i}\right\|}}{\Delta_{\alpha i}} \right. \\
&\quad \left. - \mu_{d_{P_{un}}}\right)^2
\end{aligned} \tag{7}$$

In conclusion, for Case 1, we infer from $\mu_{d_{P_{un}}}$ (Eq. 5), $\mu^*_{d_{P_{un}}}$ (Eq. 6) and $\sigma_{d_{P_{un}}}$ (Eq. 7) that $d_{P_{un}}$ has a change-dependent influence on RE (i.e., the influence is dependent on Δ_α), and the interactions between $d_{P_{un}}$ and d_I are also dependent on the change Δ_α. In specific cases, these dependencies can be experimentally further determined.

Case 2: Sensitivity analysis of F_2 keeping F_1 constant. Similarly to Case 1, we consider p_β selected levels in which $d'_{P_{un}}$ and $\overline{\Theta}$ can be set, where the Ω_β is the respective discretized input space. Then the elementary effect of $d'_{P_{un}}$ in F_2 is:

$$EE_{d'_{P_{un}}} = \frac{F_2(d_D, \overline{\Theta}) - F_2(d_D, d'_{P_{un}})}{\Delta_\beta} \tag{8}$$

where $\Delta_\beta \in \{\frac{1}{p_\beta - 1}, 1 - \frac{1}{p_\beta - 1}\}$ and $\left|\Theta\right| = \left|d'_{P_{un}}\right| \pm \Delta_\beta \geq 0$. Then the distribution $F_{d'_{P_{un}}}$ of $EE_{d'_{P_{un}}}$ is derived by randomly sampling $\overline{\Theta}$ from Ω_β.

Following the theory, we determine the mean (μ) and the standard deviation (σ) of distribution $F_{d'_{P_{un}}}$, and the mean of the absolute values (μ^*) of $\left|EE_{d'_{P_{un}}}\right|$ of the respective distribution $\left|EE_{d'_{P_{un}}}\right| \sim G_{d'_{P_{un}}}$, by applying the theory's sampling strategy. We conclude that for distributions $F_{d'_{P_{un}}}$, $G_{d'_{P_{un}}}$ which resulted from r_β samples, we have:

$$\begin{aligned}
\mu_{d'_{P_{un}}} &= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} EE_{d'_{P_{un}}i} \\
&= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{F_2(d_D, \overline{\Theta_i}) - F_2(d_D, d'_{P_{un}i})}{\Delta_{\beta i}} \\
&= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{\frac{\left|\overline{\Theta_i}\right|}{\left|d_D\right|} - \frac{\left|d'_{P_{un}i}\right|}{\left|d_D\right|}}{\Delta_{\beta i}} = \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{\frac{\left|d'_{P_{un}i}\right| \pm \Delta_{\beta i}}{\left|d_D\right|} - \frac{\left|d'_{P_{un}i}\right|}{\left|d_D\right|}}{\Delta_{\beta i}} \\
&= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{\frac{\pm\Delta_{\beta i}}{\left|d_D\right|}}{\Delta_{\beta i}} = \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{(\pm 1)_i}{\left|d_D\right|}\ (constant)
\end{aligned} \tag{9}$$

$$\begin{aligned}
\mu^*_{d'_{P_{un}}} &= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \left|EE_{d'_{P_{un}}i}\right| \\
&= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \left|\frac{F_2(d_D, \overline{\Theta_i}) - F_2(d_D, d'_{P_{un}i})}{\Delta_{\beta i}}\right| \\
&= \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \left|\frac{\frac{\pm\Delta_{\beta i}}{\left|d_D\right|}}{\Delta_{\beta i}}\right| = \frac{1}{r_\beta}\sum_{i=1}^{r_\beta} \frac{(1)_i}{\left|d_D\right|}\ (constant)
\end{aligned} \tag{10}$$

$$\begin{aligned}
\sigma^2_{d'_{P_{un}}} &= \frac{1}{r_\beta - 1}\sum_{i=1}^{r_\beta} \left(EE_{d'_{P_{un}}i} - \mu_{d'_{P_{un}}}\right)^2 \\
&= \frac{1}{r_\beta - 1}\sum_{i=1}^{r_\beta} \left(\frac{(\pm 1)_i}{\left|d_D\right|} - \mu_{d'_{P_{un}}}\right)^2\ (constant)
\end{aligned} \tag{11}$$

So, we infer from $\mu_{d'_{P_{un}}}$ (Eq. 9) and $\mu^*_{d'_{P_{un}}}$ (Eq. 10) that $d'_{P_{un}}$ has constant influence on RE, independently of the magnitude of the change (Δ_{beta}). Whereas from $\sigma_{d'_{P_{un}}}$ (Eq. 11), we infer that there are constant interactions between $d'_{P_{un}}$ and d_D independent from the magnitude of change (Δ_{beta}).

After performing the sensitivity analysis of both F_1 and F_2, we can say that all six measurements $\mu_{d_{Pun}}$ (Eq. 5), $\mu^*_{d_{Pun}}$ (Eq. 6), $\sigma_{d_{Pun}}$ (Eq. 7), $\mu_{d'_{Pun}}$ (Eq. 9), $\mu^*_{d'_{Pun}}$ (Eq. 10) and $\sigma_{d'_{Pun}}$ (Eq. 11) determined the sensitivity of F and, consequently, the behavior of the RE metric.

So, we have determined the behavior quality factor as $B = (\mu_{d_{Pun}}, \mu^*_{d_{Pun}}, \sigma_{d_{Pun}}, \mu_{d'_{Pun}}, \mu^*_{d'_{Pun}}, \sigma_{d'_{Pun}})$, for the RE.

8) STEP 8

In summary, we define:

$$RE = (\{OS, O(n), B = (\mu_{d_{Pun}}, \mu^*_{d_{Pun}}, \sigma_{d_{Pun}}, \mu_{d'_{Pun}}, \mu^*_{d'_{Pun}}, \sigma_{d'_{Pun}}), A = 1\}, F(I, D, P_{un})) \quad (12)$$

As we refer to the methodology, Eq. 12 defines a CTI quality metric and an assessment of the metric's quality characteristics in terms of objectivity, subjectivity, performance, behavior, and accuracy quality factors. In short, we observe that RE is an objective CTI quality metric calculated using partially subjective data; its calculation requires $O(n)$ time; its behavior depends on the magnitude of change over the used data, and its calculation is accurate because it introduces no bias.

B. WEIGHTED COMPLETENESS OF STRUCTURED CTI PRODUCTS

To define the weighted completeness (WC) metric of structured CTI products, we follow the methodology described in Section IV. Step 1 is skipped for brevity because this metric refers to a quality factor that characterizes the produced intelligence and specifically structured CTI products.

1) STEP 2

To identify the involved variables (members of X), we first observe that WC is related to a structured CTI product, so P_{sn} is one of the variables. Second, for the remaining variables, we need to examine the meaning of *completeness* for structured CTI products. A structured CTI product, P_{sn}, follows a well-defined structure (described by a schema $\overline{S}$) - usually a formal standard definition (e.g., STIX v2.1, IODEF). A schema ($\overline{S}$) is defined using a schema definition language (e.g., XML Schema, JSON Schema) [39], [40] and comprises metadata and object definitions ($ObjD_t, t\epsilon\mathbb{N}$) that describe the structure of P_{sn}. An $ObjD_t$ has a set of properties $\{p_1, \ldots, p_r\}$, $r\epsilon\mathbb{N}$, which have several attributes, such as minimum/maximum cardinality and occurrence. A P_{sn} comprises several $Obj_q, q\epsilon\mathbb{N}$, which are instances of $ObjD_t$. A P_{sn} is valid if it follows its schema S. Moreover, we say that a valid P_{sn} is complete if each Obj_q of P_{sn} has the maximum number of the properties defined by the respective $ObjD_t$. Hence, the set of all $ObjD_t$ is considered the second variable owing to its relation to the *completeness* quality factor because it is the basis of the determination of how complete an Obj_q is. For simplicity, we refer to this set as $S = \{ObjD_1, \ldots, ObjD_t\}$. Furthermore, for an organization C, the objects Obj_q (each being an instance of an $ObjD_t$) of a structured CTI product may be considered to have different importance (e.g., a STIX v2.1 *indicator* object may be less critical than a *malware* object). For a metric that measures the *completeness* quality factor of structured CTI products, we define vector $W = \langle w_1, w_2, \ldots, w_t \rangle$, $t = \|S\|$ & $w_t\epsilon[0, 1] \mid w_t = 0$ *by default*; w_t is a weight given by an organization to an $ObjD_t$ following its importance to that organization. In summary, we have determined the set of variables $X = \{S, W, P_{sn}\}$ as follows:

$$\begin{cases} S = \{ObjD_1, ObjD_2, \ldots., ObjD_t\}, \ t\epsilon\mathbb{N} \\ W = \langle w_1, w_2, \ldots, w_t \rangle \\ P_{sn} = \{Obj_1, Obj_2, \ldots, Obj_q\}, \end{cases}$$
$$q\epsilon\mathbb{N} \ \& \ P_{sn} \ valid \ for \ schema \ \overline{S} \quad (13)$$

2) STEP 3

To define F, we continue on the *completeness* analysis of *Step 2*, and we have:

$$\begin{aligned} &F(X) \\ &= completeness(X) \\ &= completeness(S, W, P_{sn}) \\ &= \frac{\sum(weighted)num\_of\_compl\_proper\_of\_obj\_of\_P_{sn}}{\sum(weighted)num\_of\_max\_proper\_of\_obj\_of\_P_{sn}}. \end{aligned}$$

Therefore, considering that W already has the form of a vector, we need to analyze P_{sn} and S. To use variable S in F, we need to transform it into a vector $S \rightarrow d_S$, $d_S = \langle mpr_1, \ldots, mpr_t \rangle$, $mpr_t\epsilon\mathbb{N}$, which contains the maximum number of properties mpr_t of each $ObjD_t$. To construct this vector, we examine S and count the properties by sequentially applying the rules in Table 3.

TABLE 3. Rules of counting schema properties.

Rule	Rule Definition
	For $ObjD_t$:
1	Count all its properties and set the value on variable mpr_t
2	If a choice attribute exists, which refers to x_1 properties, then subtract $x_1 - 1$ from mpr_t
3	If a minimum occurrence, x_2, and a maximum occurrence attributes x_3 exist for a property p_r, then add on x_3 to mpr_t
4	if only a maximum occurrence attribute x_3 exists, then add it to mpr_t
5	If only a minimum occurrence attribute x_2 exists, then add *alpha* to mpr_t, where *alpha* is a constant number that estimates the maximum number of properties p_r occurring in an Obj_q valid for object definition $ObjD_t$

To use P_{sn}, we need to perform the transformation $P_{sn} \rightarrow d_{P_{sn}}$, for vector $d_{P_{sn}}$ to represent the number of complete properties of Obj_q of P_{sn}. To construct $d_{P_{sn}} = \langle d_1, d_2, \ldots, d_t \rangle$, $d_t\epsilon\mathbb{N}$ we consider one dimensional zero vector, $temp = \langle 0, 0, \ldots, 0 \rangle$, of t-length. Then, for each $ObjD_t$, we count the properties of all Obj_q (instances of $ObjD_t$), named N_P_t, and the number of all Obj_q, called N_O_t, and set $temp(t) = \frac{N\_P_t}{N\_O_t}$. Finally, we set $d_{P_{sn}} = temp$. We can now define F as the fraction of the sum of the element-wise products of vector W, the element-wise division of vectors $d_{P_{sn}}$ and d_S, and the number of non-zero weights of W:

$$F(X) = F(S, W, P_{sn}) = \frac{\sum_{i=0}^{t} w_i \cdot \left[\frac{d_i}{mpr_i}\right]}{\sum_{i=0}^{t} 1 - \delta_{w_i,0}},$$
$$where\ \delta = \begin{cases} 1\ if\ w_i = 0, \\ 0\ if\ w_i \neq 0. \end{cases} \tag{14}$$

3) STEP 4

To determine *subjectivity* and *objectivity*, Γ, we observe that variable W and constant *alpha* (see Table 3), used in the computation of F, are determined by a human factor and considered subjective. Furthermore, F being the fraction of the sums of element-wise products is deterministic; as a result, F is objective. Hence, we infer that $\Gamma = OS$.

4) STEP 5

To estimate the performance of WC, we use Algorithm 2, which calculates F. We observe that F is the fraction of the sums of element-wise products by the number of non-zero elements of W. Therefore, we expect the performance of F to be $O(n)$ because we use one loop instance to calculate it. In summary, we have $P = O(n)$ for the WC metric.

Algorithm 2 *WC* Metric Algorithm

```
Require: $W = [w_1, \dots, w_t]$,$d_{P_{sn}} = [d_1, \dots., d_t]$, $d_S = [mpr_1, \dots, mpr_t]$
  $F,\ sum1, sum_2 \leftarrow 0$
  for $i$ in range length$(d_S)$ do
    $sum_1 += W[i] \cdot (d_{P_{sn}}[i]/d_S[i])$
    if $w_i > 0$ then
      $sum_2 += 1$
    end if
  end for
  $F \leftarrow sum_1/sum_2$
```

5) STEP 6

Similar to Section V-A, F does not introduce *bias* in the calculation of WC, so $A = 1$.

6) STEP 7

Similar to Section V-A, to perform a sensitivity analysis of F, we determine the variables of X for which a change is not controlled by organization C. We observe that P_{sn} is the only variable in X that can be affected by anyone outside of organization C. P_{sn} participates with $d_{P_{sn}} = \langle d_1, \dots., d_t \rangle$ in the calculation of F. We consider $\overline{Y} = \langle y_1, \dots, y_t \rangle\ y_t \epsilon \mathbb{R}$ to be the result of the change in P_{sn} that affects $d_{P_{sn}}$.

Before the sensitivity analysis of WC, we need to mention some useful observations related to d_S, $d_{P_{sn}}$, and $\overline{Y}$. By examining them, we observe that $mpr_i \epsilon [1, max(mpr_1, \dots, mpr_t)]$ and $d_i,\ y_i \leq mpr_i$.

Following the *elementary effects method* [34] for a group of factors, we consider p selected levels in which $d_{P_{sn}}$ and $\overline{Y}$ can be set, where Ω is the respective discretized input space. The elementary effect of $d_{P_{sn}}$ is:

$$EE_{d_{P_{sn}}} = \frac{F(d_S, W, \overline{Y}) - F(d_S, W, d_{P_{sn}})}{\Delta} \tag{15}$$

where $\Delta\ \epsilon\ \{\frac{1}{p-1}, 1 - \frac{1}{p-1}\}$ and $y_i = d_i \pm \Delta$. Then the distribution $F_{d_{P_{sn}}}$ of $EE_{d_{P_{sn}}}$ is then derived by randomly sampling $\overline{Y}$ from Ω.

In this case, the sensitivity measures are the mean (μ) and standard deviation (σ) of the distribution $F_{d_{P_{sn}}}$, and the mean of the absolute values (μ^*) of $\left|EE_{d_{P_{sn}}}\right|$ of the respective distribution $\left|EE_{d_{P_{sn}}}\right| \sim G_{d_{P_{sn}}}$. Here, μ assesses the influence of $d_{P_{sn}}$ in WC, μ^* assesses again the influence of $d_{P_{sn}}$ in WC simultaneously handling negative valued $EE_{d_{P_{sn}}}$, and σ reveals the total effects of the interactions between variable $d_{P_{sn}}$ and variables d_S and W.

Following the sampling approach proposed in theory, we conclude that for distributions $F_{d_{P_{sn}}}$, $G_{d_{P_{sn}}}$ derived from r samples, we have:

$$\begin{aligned}
\mu_{d_{P_{sn}}} &= \frac{1}{r}\sum_{j=1}^{r} EE_{d_{P_{sn}}j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{F(d_S, W, \overline{Y_j}) - F(d_S, W, d_{P_{sn}}j)}{\Delta_j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{\frac{\sum_{i=0}^{t} w_i \cdot \left[\frac{y_{ji}}{mpr_i}\right]}{\sum_{i=0}^{t} 1-\delta_{w_i,0}} - \frac{\sum_{i=0}^{t} w_i \cdot \left[\frac{d_{ji}}{mpr_i}\right]}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}}{\Delta_j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{\frac{\sum_{i=0}^{t} \frac{w_i}{mpr_i} \cdot \left[d_{ji} \pm \Delta_j - d_{ji}\right]}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}}{\Delta_j} = \frac{1}{r}\sum_{j=1}^{r} \frac{\frac{\sum_{i=0}^{t} \frac{w_i}{mpr_i} \cdot \left[\pm \Delta_j\right]}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}}{\Delta_j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{\sum_{i=0}^{t} \frac{(\pm 1)_j \cdot w_i}{mpr_i}}{\sum_{i=0}^{t} 1 - \delta_{w_i,0}}\ (constant)
\end{aligned} \tag{16}$$

$$\begin{aligned}
\mu^*_{d_{P_{sn}}} &= \frac{1}{r}\sum_{j=1}^{r} \left|EE_{d_{P_{sn}}j}\right| = \frac{1}{r}\sum_{j=1}^{r} \frac{\left|\frac{\sum_{i=0}^{t} \frac{w_i}{mpr_i} \cdot \left[\pm \Delta_j\right]}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}\right|}{\Delta_j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{\frac{\sum_{i=0}^{t} \frac{w_i}{mpr_i} \cdot \Delta_j}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}}{\Delta_j} = \frac{1}{r}\sum_{j=1}^{r} \frac{\frac{\Delta_j \sum_{i=0}^{t} \frac{w_i}{mpr_i}}{\sum_{i=0}^{t} 1-\delta_{w_i,0}}}{\Delta_j} \\
&= \frac{1}{r}\sum_{j=1}^{r} \frac{\sum_{i=0}^{t} \frac{w_i}{mpr_i}}{\sum_{i=0}^{t} 1 - \delta_{w_i,0}}\ (constant)
\end{aligned} \tag{17}$$

$$\begin{aligned}
\sigma^2_{d_{P_{sn}}} &= \frac{1}{r-1}\sum_{j=1}^{r} \Big(EE_{d_{P_{sn}}j} - \mu_{d_{P_{sn}}}\Big)^2 \\
&= \frac{1}{r-1}\sum_{j=1}^{r} \Big(\frac{\sum_{i=0}^{t} \frac{(\pm 1)_j \cdot w_i}{mpr_i}}{\sum_{i=0}^{t} 1 - \delta_{w_i,0}} - \mu_{d_{P_{sn}}}\Big)^2\ (constant)
\end{aligned} \tag{18}$$

In conclusion, we can infer from $\mu_{d_{P_{sn}}}$ (Eq. 16) and $\mu^*_{d_{P_{sn}}}$ (Eq. 17) that $d_{P_{sn}}$ has a constant influence on WC independent of the magnitude of the change (Δ). Whereas from $\sigma_{d_{P_{sn}}}$ (Eq. 18), we infer that there are interactions between variable $d_{P_{sn}}$ and variables d_S and W whose level should be experimentally determined in each case, but are again independent of the magnitude of change (Δ).

After performing the sensitivity analysis, we determine the behavior quality factor as $B = (\mu_{d_{P_{sn}}}, \mu^*_{d_{P_{sn}}}, \sigma_{d_{P_{sn}}})$, for WC.

7) STEP 8

In summary, we define:

$$WC = (\{OS, O(n), B = (\mu_{d_{P_{sn}}}, \mu^*_{d_{P_{sn}}}, \sigma_{d_{P_{sn}}}), A = 1\}, F(S, W, P_{sn})) \quad (19)$$

Similarly to Eq. 12, Eq. 19 defines a CTI quality metric (WC) and provides an assessment of the WC metric's quality characteristics. Again, we observe that WC is an objective metric calculated on partially subjective data; its calculation requires O(n) time. WC's behavior is independent of the magnitude of change over the data used, and its calculation is accurate because it does not introduce bias. Consequently, even if RE and WC are defined for different types of CTI products, we can deduce that RE is more sensitive to changes (i.e., changes in its input data) than WC.

TABLE 4. Datasets statistics.

Dataset	Description	Value
I	Application CPEs	818333
	Operating System CPEs	84954
	Hardware CPEs	2239
	Total CPEs	**945526**
D	Industry domains	26
S	STIX V2.1 Schema sdos objects	19
	STIX V2.1 Schema observable objects	18
	STIX V2.1 Schema common objects	18
	STIX V2.1 Schema sros objects	2
P_u	Alerts of ACSC	97
	Alerts of BGD e-GOV CIRT	589
	Alerts of CISA	473
	Alerts of JPCERT/CC	547
	Total alert files	**1706**
	Average words per alert	*6956*
	Reports of ACSC	10
	Reports of CISA	113
	Total report files	**123**
	Average words per report	*2478*
P_s	STIX V2.1 Bundle objects	50
	STIX V2.1 Indicator objects	67307
	STIX V2.1 Report objects	50
	STIX V2.1 Identity objects	50
	STIX V2.1 Vulnerability objects	4
	STIX V2.1 Threat Actor objects	4
	STIX V2.1 Observ. URL objects	67
	STIX V2.1 Observ. Domain Name obj.	379
	STIX V2.1 Observ. IPv4 Address obj.	66173
	STIX V2.1 Observ. Email Message obj.	17
	STIX V2.1 Observable File objects	613
	Total STIX V2.1 objects	**134714**

VI. COMPARISON OF CTI QUALITY METRICS

A. DATASETS DEVELOPMENT

To demonstrate and experimentally measure the developed CTI quality metrics, we created the datasets I, D, S, P_u, P_s. The dataset I was constructed by manipulating the CPE v.3 dictionary. D was constructed by extracting the data from the DIT full taxonomy. S was constructed by using the STIX v2.1 Schema. Finally, because our research in public repositories indicates a lack of benchmark datasets related to unstructured and structured CTI Products, we have constructed P_u and P_s. Specifically, we constructed the P_u dataset by collecting alerts from the Cybersecurity & Infrastructure Security Agency (CISA) [41], the Japan Computer Emergency Response Team Coordination Center (JPCERT/CC) [42], the Bangladesh Government's e-Government Computer Incident Response Team (BGD e-GOV CIRT) [43], and the Australian Cyber Security Centre (ACSC) [44]. We constructed P_s by collecting STIX v2.1 documents from AlienVault Open Threat Exchange (OTX) [45]. Table 4 presents the statistics of the created datasets (see also Appendix).

B. EXPERIMENTAL RESULTS

1) RELEVANCE (RE) EXPERIMENTAL RESULTS

To conduct the various measurements using the RE metric, we consider ten organizations $C_1, \ldots, C_{10}$ and split P_u dataset into ten randomly created subsets of data $P_{u1}, \ldots, P_{u2}$ with an equal number of alerts and reports files. Moreover, we create an I_i for each organization C_i by randomly selecting $i * 10000$ thousand CPEs. We also create a D_i for each C_i by randomly selecting i industry domains. Figure 1 shows the RE experimental results.

In Figure 1(a), we present the calculation of the average RE metric value of each P_{ui} per organization, demonstrating how RE can be used to evaluate the relevance of an aggregation of unstructured CTI products with an organization. In Figure 1(b), we use the RE to compare the relevance of two unstructured CTI products within each organization. Figure 1(b) also shows an organization's ability to select the most relevant unstructured CTI products using RE. Finally, in Figure 1(c), to measure the performance of RE, we define for each organization an $artificial\_product_i = i * 10 * P_{si}$. We have measured the RE calculation time for each artificial product and have approximately verified that the performance of the RE calculation algorithm is $P = O(n)$.

2) WEIGHTED COMPLETENESS (WC) EXPERIMENTAL RESULTS

To conduct the various measurements of the WC metric, we consider ten organizations $C_1, \ldots, C_{10}$ and we split P_s into ten randomly created subsets of data $P_{s1}, \ldots, P_{s10}$ with an equal number of items of different STIX v2.1 objects. We randomly create a W_i for each organization C_i. Figure 2 shows the WC experimental results.

In Figure 2(a), we present the calculation of the average WC metric value of each object type of STIX v2.1 in the ten organizations. We observe each organization's different WC

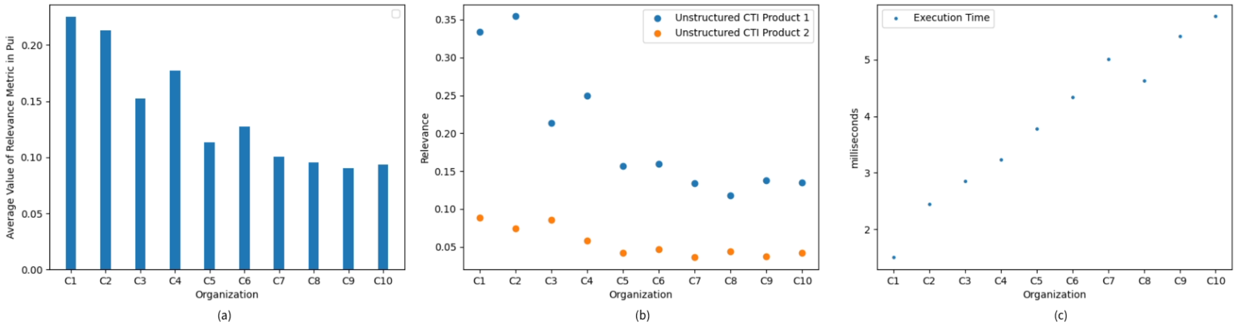

FIGURE 1. Experimental results for the *RE* metric: (a) Average *RE* of each P_{ui}, (b) Relevance comparison of two CTI products, and (c) Performance measurement of *RE*.

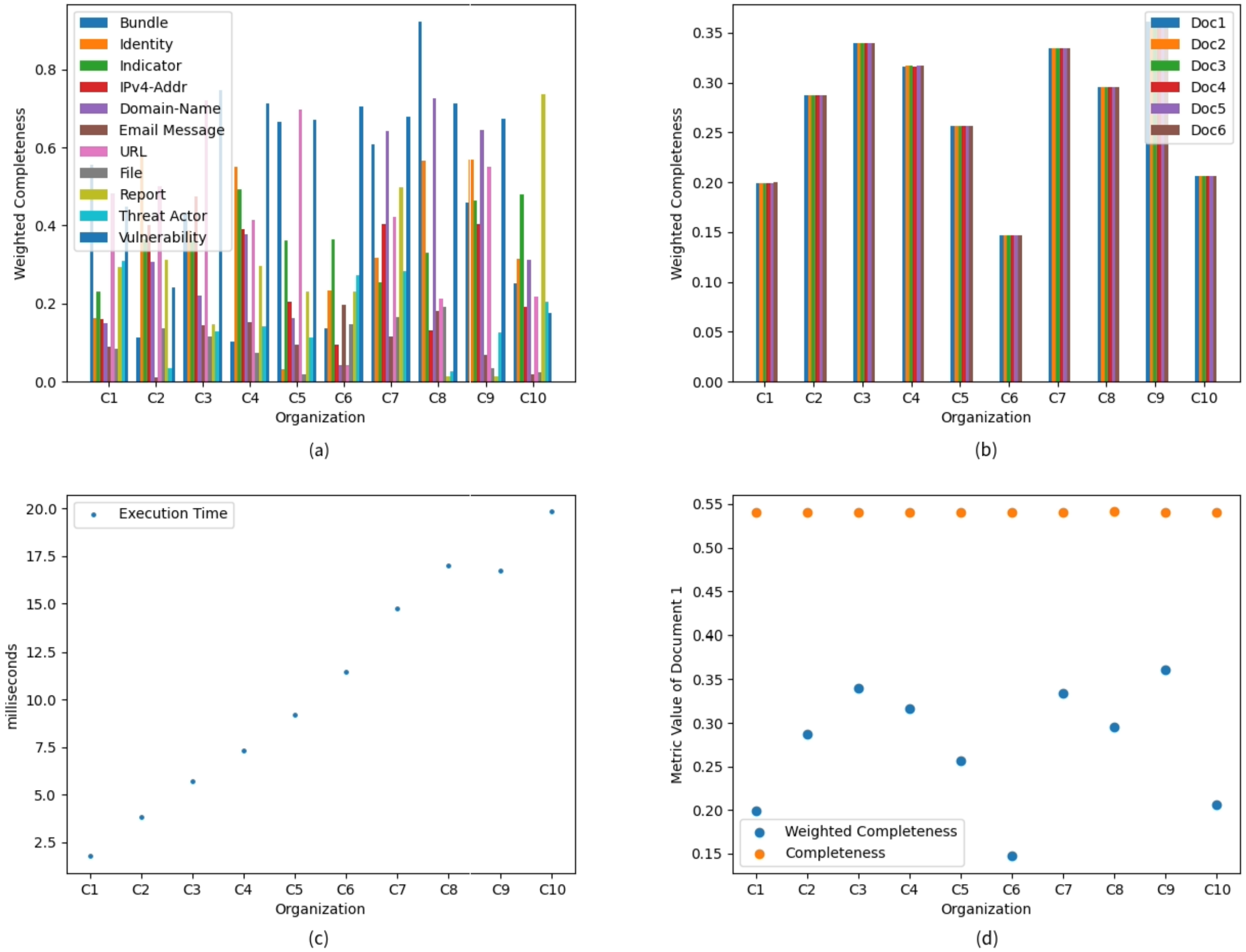

FIGURE 2. Experimental results for the *WC* Metric: (a) Average *WC* of each object type, (b) organization's *WC* of six artificial structured documents, (c) Performance measurement of *WC*, and (d) Comparison of *WS* with the completeness metric proposed in [9].

metric values for the same object type. This highlights the ability of an organization to tune *WC* based on the perceived importance of each object type. In Figure 2(b), we create six artificially structured documents (Doc1,..., Doc6) for each organization by splitting each $P_{s1}, \ldots, P_{s10}$ into six equal sets of objects, in which we calculate *WC*. We again observe the role of weights in the estimation of *WC*; however, the critical part is that the metric converges when it is calculated in structured CTI products (i.e., Doc1,..., Doc6) with a large number of similar objects, independently of the weighted completeness of each object. This characteristic of *WC* can be used in the context of Big Data because it seems that by measuring the *WC* of a representative sample of a large dataset, we can estimate the overall *WC* of it. In Figure 2(c), we define $artificial\_product_i = i{*}P_{si}$. We have measured *WC* calculation execution time for each artificial product and have verified that the performance of calculating *WC* is $P = O(n)$.

For completeness, we recall that few metrics have been proposed in the literature (see Section II); however, in Figure 2(d), we use the previously described Doc1 to compare *WC* with the proposed completeness metric according to Schaberreiter et al. [9]. For the latter metric, we consider the total number of properties of an object type as the *"world view"* and the number of properties of an object as the *"coverage"* of this *"world view"*. We observe that *WC* yields lower values than the metric of Schaberreiter et al. [9]. This again highlights an organization's ability to tune *WC* based on the perceived object type importance.

VII. DISCUSSION

In the context of a CTI community, since the value of an "information" piece differs for each community member, we should expect that community members may employ (within their organizations) different quality metrics to evaluate the CTI data and sources. However, sharing CTI data and source evaluations within a community is challenging, although it enhances collaboration between community members. To address such a sharing requirement, a minimum set of common CTI quality metrics should be established (acceptable to all members). The proposed methodology integrates the quality characteristics of a metric into its definition, allowing the comparison of alternative metrics. By utilizing the comparison capability provided by the proposed methodology, the community can assess and accept a set of CTI quality metrics whose quality characteristics are included in the respective definitions.

Furthermore, the proposed methodology offers community members the flexibility to adapt (e.g., by adding new variables on X) community-defined CTI quality metrics to particular requirements of their organizations. This flexibility contributes positively to member decisions, empowering the establishment of a set of CTI quality metrics (e.g., an assessment baseline), for exchanging evaluations. For example, the Common Vulnerability Scoring System (CVSS) metric is used in vulnerability management to evaluate the severity of a particular vulnerability. CVSS consists of the CVSS Base score (the same for all members), the CVSS Temporal score (which tracks the severity change of a vulnerability over its lifetime), and the CVSS Environmental score (which represents the effect of the security requirements of a specific organization on the severity of a vulnerability). CVSS Base can be considered as the metric used by all members of the vulnerability management community, while any particular organization member can modify the CVSS Base score by calculating and using its own CVSS Environmental score.

In Section IV, we analyzed the quality factors of CTI quality metrics. To the best of our knowledge, this is the first time that the behavior and accuracy quality factors of CTI quality metrics are discussed. This study commences the exploration of the behavior quality factor and the methods needed to estimate it. Moreover, further research on developing methods determining the accuracy of CTI quality metrics for which a non-deterministic algorithm could be employed for calculating F seems to be necessary.

Finally, in this study, we proposed two CTI quality metrics focusing on CTI products to demonstrate the applicability of the proposed methodology. Therefore, we recognize the importance of developing CTI quality metrics for CTI sources evaluation in the context of future research.

VIII. CONCLUSION

In this study, we have analyzed CTI evaluation needs by proposing an approach for quantitative measurement of CTI quality. We have set three research questions (Q1-Q3) to achieve this goal in Section I. To answer those research questions, we have analyzed the current state of the CTI quality factors and their respective metrics, proposed a CTI quality metrics development and assessment methodology, and applied this methodology to develop two CTI quality metrics.

Specifically, we have analyzed their characteristics to identify the quality factors that better evaluate raw CTI data, CTI products, and CTI sources. We also have determined the areas of CTI in which CTI quality factors can be categorized based on the bibliography. The combined result of this effort was to determine the relation of CTI quality factors with the CTI data and CTI sources, thereby setting the base for the proposed methodology.

Next, to address the definition of metrics of CTI quality factors, we have proposed a systematic methodology that provides a systematic way to develop CTI quality metrics for CTI data and CTI sources, defines and determines the quality characteristics of a metric itself ($Q = \{\Gamma, P, B, A\}$), and produces easily comparable metrics since a metric's quality characteristics measurements are included in its definition.

To prove the applicability of the proposed methodology, we have developed two metrics: one for unstructured CTI products (RE), and one for structured CTI products (WC), with the following characteristics:

- The application of text similarity as a metric for comparing an organization's characteristics (I, D) with the unstructured CTI product (P_{un}).
- The tailoring of completeness metric (WC) of structured CTI products in the needs of an organization by introducing a weighted approach.
- The introduction of a standard-independent approach to defining the *WC* metric.

We have conducted several experiments to demonstrate the efficiency of these two metrics. As part of those experiments and due to the lack of publicly available datasets, we have created two datasets, one from unstructured CTI products and one from structured CTI products, by collecting them from open sources. These two datasets are available for future research under the GNU v3 General Public License.

This study addresses the topic of CTI quality by introducing a systematic methodology for developing CTI quality metrics. To the best of our knowledge, this is one of the first studies that follows such an approach in the field of CTI. It also opens a path for future research to improve the integration of CTI in existing cybersecurity operations. Finally, we consider the integration of CTI quality metrics into existing or new CTI systems as an intriguing research direction because CTI quality metrics can enhance CTI product utilization and exploration by cybersecurity specialists.

APPENDIX

ONLINE RESOURCES

The source code and datasets of the experiments are available under a GNUv3 General Public License in the respective repository: https://github.com/geosakel77/s2

REFERENCES

[1] G. Sakellariou, P. Fouliras, I. Mavridis, and P. Sarigiannidis, "A reference model for cyber threat intelligence (CTI) systems," *Electronics*, vol. 11, no. 9, p. 1401, Apr. 2022.

[2] S. Samtani, M. Abate, V. Benjamin, and W. Li, *Cybersecurity as an Industry: A Cyber Threat Intelligence Perspective*. Cham, Switzerland: Springer, 2020, pp. 135–154.

[3] T. D. Wagner, K. Mahbub, E. Palomar, and A. E. Abdallah, "Cyber threat intelligence sharing: Survey and research directions," *Comput. Secur.*, vol. 87, Nov. 2019, Art. no. 101589.

[4] Ernst & Young Global Limited, "Cyber threat intelligence—How to get ahead of cybercrime," *Insights Goverance, Risk Compliance*, vol. 1, no. 1, pp. 1–16, 2014.

[5] C. Beard, S. Brown, A. Dulaunou, J. Ginn, and P. Stipraro, "Exploring the opportunities and limitations of current threat intelligence platforms," ENISA, Athens, Tech. Rep. 1, 2017.

[6] A. Zibak, C. Sauerwein, and A. C. Simpson, "Threat intelligence quality dimensions for research and practice," *Digit. Threats, Res. Pract.*, vol. 3, no. 4, pp. 1–22, Dec. 2022.

[7] D. Schlette, F. Böhm, M. Caselli, and G. Pernul, "Measuring and visualizing cyber threat intelligence quality," *Int. J. Inf. Secur.*, vol. 20, no. 1, pp. 21–38, Feb. 2021.

[8] L. Qiang, J. Zhengwei, Y. Zeming, L. Baoxu, W. Xin, and Z. Yunan, "A quality evaluation method of cyber threat intelligence in user perspective," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun., 12th IEEE Int. Conf. Big Data Sci. Eng. (Trustcom/BigDataSE)*. New York, NY, USA: Institute of Electrical and Electronics Engineers, Sep. 2018, pp. 269–276.

[9] 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 *Proc. ACM Int. Conf. Proc. Ser.*, New York, NY, USA: Association for Computing Machinery, Aug. 2019, pp. 1–10.

[10] R. Meier, C. Scherrer, D. Gugelmann, V. Lenders, and L. Vanbever, "FeedRank: A tamper- resistant method for the ranking of cyber threat intelligence feeds," in *Proc. 10th Int. Conf. Cyber Conflict (CyCon)*, May 2018, pp. 321–344.

[11] N. Dalkey, O. Helmer, and O. Helmer, "An experimental application of the delphi method to the use of experts," *Manage. Sci.*, vol. 9, no. 3, pp. 458–467, Apr. 1963, doi: 10.1287/mnsc.9.3.458.

[12] C. Martins and I. Medeiros, "Generating quality threat intelligence leveraging OSINT and a cyber threat unified taxonomy," *ACM Trans. Privacy Secur.*, vol. 25, no. 3, pp. 1–39, Aug. 2022.

[13] A. Pinto and A. Sieira, "Data-driven threat intelligence: Useful methods and measurements for handling indicators," FIRST, North Carolina, Tech. Rep. 1, 2015.

[14] P. Pawliński, P. Kijewski, and A. D. Kompanek, "Towards a methodology for evaluating threat intelligence feeds," 2016.

[15] S. Zhang, P. Chen, G. Bai, S. Wang, M. Zhang, S. Li, and C. Zhao, "An automatic assessment method of cyber threat intelligence combined with ATT&CK matrix," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–12, Aug. 2022.

[16] G. Grispos, W. B. Glisson, and T. Storer, "How good is your data? Investigating the quality of data generated during security incident response investigations," in *Proc. 52nd Hawaii Int. Conf. Syst. Sci.* Scholar Space Hawaii International, Apr. 2019, p. 10.

[17] H. Dalziel, *How to Define and Build an Effective Cyber Threat Intelligence Capability*. London, U.K.: Syngress, an imprint of Elsevier, 2015.

[18] J. A. Friedman and R. Zeckhauser, "Assessing uncertainty in intelligence," *Intell. Nat. Secur.*, vol. 27, no. 6, pp. 824–847, Dec. 2012.

[19] S. E. Jasper, "U.S. cyber threat intelligence sharing frameworks," *Int. J. Intell. CounterIntelligence*, vol. 30, no. 1, pp. 53–65, Jan. 2017.

[20] D. Schlette, F. Böhm, M. Caselli, and G. Pernul, "Measuring and visualizing cyber threat intelligence quality," *Int. J. Inf. Secur.*, vol. 20, no. 1, pp. 21–38, Feb. 2021.

[21] C. Sillaber, C. Sauerwein, A. Mussmann, and R. Breu, "Data quality challenges and future research directions in threat intelligence sharing practice," in *Proc. ACM Workshop Inf. Sharing Collaborative Secur.*, New York, NY, USA, 2016, pp. 65–70.

[22] W. Tounsi and H. Rais, "A survey on technical threat intelligence in the age of sophisticated cyber attacks," *Comput. Secur.*, vol. 72, pp. 212–233, Jan. 2018.

[23] V. Mavroeidis and S. Bromander, "Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence," in *Proc. Eur. Intell. Secur. Informat. Conf. (EISIC)*, Athens, Greece, Sep. 2017, pp. 91–98.

[24] Y. Cheng, J. Deng, J. Li, S. A. DeLoach, A. Singhal, and X. Ou, "Metrics of security," in *Cyber Defense and Situational Awareness*, A. Kott, C. Wang, and R. Erbacher, Eds. Cham, Switzerland: Springer, 2014, pp. 263–295.

[25] C. Bolstad and H. Cuevas, "Integrating situation awareness assessment into test and evaluation," *ITEA J.*, vol. 31, pp. 240–246, Jan. 2010.

[26] A. Evesti, T. Kanstren, and T. Frantti, "Cybersecurity situational awareness taxonomy," in *Proc. Int. Conf. Cyber Situational Awareness, Data Anal. Assessment, Cyber SA*, New York, NY, USA: Institute of Electrical and Electronics Engineers, Oct. 2017, pp. 1–8.

[27] C. Onwubiko, "Understanding cyber situation awareness," *Int. J. Cyber Situational Awareness*, vol. 1, no. 1, pp. 11–30, Dec. 2016.

[28] A. E. Fridman, *The Quality of Measurements: A Metrological Reference*, vol. 1. New York, NY, USA: Springer, Jul. 2012.

[29] F. A. Muckler and S. A. Seven, "Selecting performance measures: 'Objective' versus 'subjective' measurement," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 34, pp. 441–455, Nov. 2016, doi: 10.1177/001872089203400406.

[30] J. T. Buchanan, E. J. Henig, and M. I. Henig, "Objectivity and subjectivity in the decision making process," *Ann. Oper. Res.*, vol. 80, pp. 333–345, 1998.

[31] J. M. Rothstein, "On defining subjective and objective measurements," *Phys. Therapy*, vol. 69, no. 7, pp. 577–579, Jul. 1989.

[32] T. Mahlangu, S. January, T. Mashiane, M. Dlamini, S. Ngobeni, and N. Ruxwana, "Data poisoning: Achilles heel of cyber threat intelligence systems," in *Proc. 14th Int. Conf. Cyber Warfare Secur. (ICCWS)*, Cape Town, South Africa, Stellenbosch, South Africa: Stellenbosch Univ., Feb. 2019, pp. 220–230.

[33] A. Saltelli, "Sensitivity analysis for importance assessment," *Risk Anal.*, vol. 22, no. 3, pp. 579–590, Jun. 2002.

[34] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis: The Primer*. New York, NY, USA: Wiley, Jan. 2008.

[35] *Official Common Platform Enumeration (CPE) Dictionary*, CFP NIST, 2021.

[36] *Domain Industry Taxonomy*, RRDG, 2022.

[37] A. Gakhov, *Probabilistic Data Structures and Algorithms for Big Data Applications*, 1st ed. Norderstedt, Germany: BoD-Books on Demand, 2022.

[38] K. Karthikeyan, "Time complexity for document similarity measures," 2022.

[39] *W3C XML Schema*, W3C, 2022.

[40] *JSON Schema | The home of JSON Schema*, JSON Schema Organization, 2022.

[41] *Alerts | Cybersecurity & Infrastructure Security Agency*, CISA, 2022.
[42] *Security Alerts | Japan Computer Emergency Response Team Coordination Center*, JPCERT/CC, 2022.
[43] *Security Advisories & Alerts Bangladesh Government's e-Government Computer Incident Response Team*, BGD e-GOV CIRT, 2022.
[44] *Australian Cyber Security Centre*, ACSC, 2022.
[45] *AlienVault—Open Threat Exchange*, OTX, 2022.

GEORGIOS SAKELLARIOU received the B.S. degree from Hellenic Army Academy, the M.Sc. degree in applied informatics from the University of Macedonian (UoM) and the M.Sc. degree in knowledge, data, and software technologies from the Aristotle University of Thessaloniki (AUTH). He is currently pursuing the Ph.D. degree with UoM. He is a Vulnerability Engineer at the international organization. He is also a member of the Multimedia, Security and Networking (MSN) Laboratory, UoM. His research interests include the design and integration of cyber threat intelligence in cyber defense mechanisms and particularly the collection and analysis of cyber threat-related data and the quality of CTI data and sources.

PANAGIOTIS FOULIRAS received the B.Sc. degree in physics from the Aristotle University of Thessaloniki, Greece, and the M.Sc. and Ph.D. degrees in computer science from the University of London, U.K. (QMW). He is currently a permanent Assistant Professor with the Department of Applied Informatics, University of Macedonia, Thessaloniki. He has participated in several national and European-funded (H2020) research projects and published articles in many international journals. His research interests include computer networks and network security, blockchain, and system evaluation methods.

IOANNIS MAVRIDIS received the Diploma degree in computer engineering and the M.A. degree from the University of Patras, Greece, and the Ph.D. degree in information systems security from the Aristotle University of Thessaloniki, Greece. He is currently a Professor of information security with the Department of Applied Informatics, University of Macedonia, Greece. He is currently the Director of the Multimedia, Security and Networking (MSN) Laboratory. He has published more than 100 articles in journals and conferences. He is the author or coauthor of three books on information security. He has participated as a principal investigator and a researcher of several international and nationally-funded research and development projects. His current research interests include cybersecurity education on risk management, access control, cyber threat intelligence, digital forensics, and security economics. He serves as an Area Editor for *Journal of Cyber Security* and *Array* (Elsevier).