

Article **A Systematic Analysis of Security Metrics for Industrial Cyber–Physical Systems**

Giacomo Gori [,](https://orcid.org/0009-0002-3218-6703) Lorenzo Rinieri [,](https://orcid.org/0000-0003-1767-7256) Andrea Melis [,](https://orcid.org/0000-0002-0101-2551) Amir Al Sadi [,](https://orcid.org/0000-0002-5560-9871) Franco Callegati and Marco Prandini [*](https://orcid.org/0000-0002-3962-5513)

Department of Computer Science and Engineering (DISI), Alma Mater Studiorum, Università di Bologna, 40136 Bologna, Italy; g.gori@unibo.it (G.G.); lorenzo.rinieri@unibo.it (L.R.); a.melis@unibo.it (A.M.); amir.alsadi@unibo.it (A.A.S.); franco.callegati@unibo.it (F.C.)

***** Correspondence: marco.prandini@unibo.it

Abstract: Nowadays, as the cyber-threat landscape is evolving and digital assets are proliferating and becoming more and more interconnected with the internet and heterogeneous devices, it is fundamental to be able to obtain a sensible measure of the security of devices, networks, and systems. Industrial cyber–physical systems (ICPSs), in particular, can be exposed to high operational risks that entail damage to revenues, assets, and even people. A way to overcome the open question of measuring security is with the use of security metrics. With metrics it is possible to rely on proven indicators that benchmark systems, identify vulnerabilities, and show practical data to assess the risk. However, security metrics are often proposed with specific contexts in mind, and a set of them specifically crafted for ICPSs is not explicitly available in the literature. For this reason, in this work, we analyze the current state of the art in the selection of security metrics and we propose a systematic methodology to gather, filter, and validate security metrics. Then, we apply the procedure to the ICPS domain, gathering almost 300 metrics from the literature, analyzing the domain to identify the properties useful to filter the metrics, and applying a validation framework to assess the validity of the filtered metrics, obtaining a final set capable of measuring the security of ICPSs from different perspectives.

Keywords: cyber–physical systems; cybersecurity; industrial cyber–physical systems; industrial networks; security metrics

1. Introduction

In recent years, as technology advances, the growing importance and criticality of cybersecurity have become undeniable. The escalating cyber-threat landscape leverages ubiquitous digital technology and global interconnectivity, raising concerns for people's well-being, damaging the economy, and putting national security at risk; enacting robust cybersecurity measures is a priority.

A major issue regarding this topic is understanding what security policies to implement and which of them are working efficiently: to find the best decisions, it is important to perform activities such as risk assessment and performance evaluation and periodically generate reports, intending to reach a continuous improvement in decision making. Security metrics play an essential role in this process. They are quantitative or qualitative measurements, produced over time, and used to evaluate and assess various aspects of security in computer systems. Nowadays, security metrics are widely used in organizations to deal with risk management, compliance, and regulation, e.g., ISO 27004 [\[1\]](#page-14-0) and NIST 800-55 [\[2\]](#page-14-1). Most of the standards refer in particular to policies for organizations and procedures rather than quantifying the security of systems themselves. As of now, many initiatives, such as the European Cybersecurity Act [\[3\]](#page-14-2), strive to reach a standardized set of security metrics for the evaluation of systems as an enabler of comparative analysis between different systems: this opens new challenges regarding pinpointing the properties and features to factor in the different assessment types.

Citation: Gori, G.; Rinieri, L.; Melis, A.; Al Sadi, A.; Callegati, F.; Prandini, M. A Systematic Analysis of Security Metrics for Industrial Cyber–Physical Systems. *Electronics* **2024**, *13*, 1208. [https://doi.org/](https://doi.org/10.3390/electronics13071208) [10.3390/electronics13071208](https://doi.org/10.3390/electronics13071208)

Academic Editors: Georgios Spathoulas, Georgios Kavallieratos and Marios Anagnostopoulos

Received: 19 February 2024 Revised: 18 March 2024 Accepted: 20 March 2024 Published: 25 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license [\(https://](https://creativecommons.org/licenses/by/4.0/) [creativecommons.org/licenses/by/](https://creativecommons.org/licenses/by/4.0/) $4.0/$).

To formulate a standard set of security metrics we need to tailor them to the specific context, define a way to collect them, and then validate the effectiveness, efficiency, and

feasibility of each of them. To better understand the effectiveness of a validated set of security metrics and show tangible preliminary results, we chose a real use-case scenario, industrial cyber–physical systems (ICPSs), in which to perform a systematic approach to collect, filter, and validate metrics that can be used in practice.

Given their complexity, ICPSs necessitate an increasing reliance on automation involving the intelligent and self-governing operation of individual subsystems and their coordination at the infrastructure level. Algorithms play a crucial role in activating and configuring components, as well as facilitating their connectivity and interaction with the physical world. In making decisions, functional requirements, system and network parameters, and sensor measurements are taken into consideration, as security properties should be. Here, security metrics find their essential role, to drive those configurations towards a more secure state: in this scenario, it is not enough to evaluate every single component separately but there is the need to evaluate the overall system, so the metrics should take into account the communication and collaboration that physical devices and logical components establish with each other. Therefore, we can identify the need to use only the metrics that are applicable and meaningful for ICPSs. In this domain, threats include spoofing identity, tampering with data, repudiation of origin, information disclosure, elevation of privilege, and denial of service (DoS).

For this reason, in this work, we analyze the current state of the art in the selection of security metrics and we propose a methodology to gather, filter, and validate security metrics. Then, we apply the procedure to the ICPS domain, gathering 291 metrics from the literature, analyzing the domain to identify the properties useful to filter the metrics, and applying a validation framework to assess the validity of the filtered metrics, obtaining a final set capable of measuring security from different perspectives.

This paper proceeds hereinafter with an illustration of the state of the art in the literature about security metrics, with particular reference to their application in ICPSs. The third section describes the selection criteria we followed to collect a set of metrics available in the literature based on a variation of the classification and selection strategy. We applied the systematic criteria as detailed in the fourth section, which presents the phases, based on the validation and filtering approaches, that we followed to derive a set of security metrics for ICPSs. Finally, the fifth section discusses the outcomes of our process and its limitations, before the sixth and final section presents the conclusions of this study.

2. State of the Art

Defining standard security metrics is not straightforward, and this is confirmed by Philippou et al. [\[4\]](#page-14-3), who criticize the works we mentioned in the Introduction [\[1](#page-14-0)[,2\]](#page-14-1). They argue that there is a lack of proper contextualization and alignment with business objectives and, to answer this, they suggest a new strategy. Hence, they demonstrate how the proposed method offers more precise and suitable outcomes, albeit with the drawback of demanding significant effort to establish a clear and traceable connection between metrics and business objectives.

As contextualization depends on the considered domain, several studies have extensively delved into specific areas. For instance, Wang et al. [\[5\]](#page-14-4) concentrate on network security metrics and assess the advantages and disadvantages of each metric. On the other hand, Longueira-Romero et al. [\[6\]](#page-14-5) undertake a rigorous filtering process to identify metrics suitable for embedded systems applications. Initially, more than 200 metrics were considered, and from this pool, 169 metrics were selected for evaluation utilizing criteria such as SMART [\[7\]](#page-14-6) and PRAGMATIC [\[8\]](#page-14-7), along with characteristics drawn from the work of Savola et al. [\[9\]](#page-14-8). The focus was primarily on assessing the comparability, cost effectiveness, measurability, repeatability, and reproducibility of each metric.

On the other hand, some research has adopted a survey approach, such as Pendleton et al. [\[10\]](#page-14-9), which compares various proposals in system security. This survey measures the effectiveness of security metrics in terms of vulnerabilities, severity of attacks, and defense mechanisms' strength. The findings underscore significant gaps between the research outcomes available and the desired properties of metrics. Although specific sub-fields, such as security conformance metrics for managing industrial automation control systems, provide clear definitions of desirable metric properties, as outlined in Hauet's commentary of the ISA99/IEC62443 standard [\[11\]](#page-14-10), a need for greater clarity in defining such properties still remains.

The first rough selection of suitable metrics is arguably the most complex and tricky phase since there are no standards to perform it. In the literature, there are some proposals on how to address this task, such as:

- Classification and selection: Used by Sultan et al. [\[12\]](#page-14-11) and Morrison et al. [\[13\]](#page-14-12), defines a classification of the metrics coherent with the domain, then selects them to cover all the security aspects that are required.
- Automatic generation: Used by Ani et al. [\[14\]](#page-15-0), i.e., a framework that generates specific security metrics after a preliminary study that analyzes the context and the security objectives of that field.
- Multivocal literature review (MLR): Used by Fernandez et al. [\[15\]](#page-15-1), that consists of exploring the academic and gray literature, using the snowballing process and filtering them with a multi-step approach.

Context significantly affects the impacts of security metrics. Even similar applications such as traditional biometric systems and wearable biometric systems can have different be-havior, as shown in [\[16\]](#page-15-2), for example, due to variations in relevant threats and vulnerabilities.

To assist in identifying suitable metrics, it is important to follow a proper breakdown of metrics into their categories of technical domains. The classification can follow various criteria, but a prevalent approach in the literature, as also referenced in [\[10\]](#page-14-9), classifies technical metrics into four distinct types:

- Defense metrics: These metrics assess the strength and effort required to implement defense mechanisms within a system. They encompass the evaluation of preventive, reactive, and proactive defenses, as explored in [\[17\]](#page-15-3).
- Vulnerability metrics: These metrics gauge system vulnerabilities, encompassing user vulnerabilities, interface-induced vulnerabilities, and software vulnerabilities. Examples include password vulnerabilities, attack surface [\[18\]](#page-15-4), and software vulnerabilities, as documented by the Common Vulnerability Scoring System (CVSS) [\(https://nvd.nist.gov/vuln-metrics/cvss,](https://nvd.nist.gov/vuln-metrics/cvss) accessed on 19 March 2024).
- Attack metrics: These metrics quantify the strength of performed cyberattacks. Unlike the previous categories that assess the security level via configuration and device analysis, attack metrics concentrate on measuring and analyzing cyberattacks and threats. They are crucial for risk assessment, evaluating the success of security measures, and guiding resource allocation. Examples include network bandwidth used by a botnet for launching denial-of-service attacks, the occurrence of obfuscation in malware samples, or the runtime complexity of packers measured in the number of layers or granularity [\[19\]](#page-15-5).
- Situation metrics: Focusing on the security state of a system, situation metrics are timedependent and dynamically evolve based on attack–defense interactions. Examples include metrics based on the frequency of security incidents or those related to investment in security improvement [\[20\]](#page-15-6). They are further categorized into data-driven metrics, such as the network maliciousness metric [\[21\]](#page-15-7), and model-driven metrics, such as the fraction of compromised computers.

The evaluation of a metric is outlined in a study conducted by Ahmed et al. [\[22\]](#page-15-8), which considers the following aspects:

- The measurability of properties that should be consistently accessible.
- The feasibility and potential for automated data collection, taking into account associated costs.
- The methodology for quantifying the metric, such as using cardinal numbers or percentages.

The establishment of units for measurement.

According to Savola et al. [\[9\]](#page-14-8), good metrics should give a simple answer, such as a score. An example of this approach is the CVSS, a published standard and open framework that "provides a way to capture the principal characteristics of a vulnerability and produce a numerical score reflecting its severity" [\(https://www.first.org/cvss/,](https://www.first.org/cvss/) accessed on 19 March 2024), with a calculator available on the website. The score is a result of the combination of various factors that depends on qualities intrinsic to a vulnerability, e.g., the skills required to exploit it; qualities that can change over the lifetime of a vulnerability, e.g., if patches are available; or even qualities that depend on the context, such as impact on physical and financial assets.

Moreover, Yee et al. [\[23\]](#page-15-9) propose conditions that can be used to design and test security metrics' soundness, improving a method they previously suggested. Furthermore, they demonstrate that the aggregation of sound security metrics results in a new metric that also respects the conditions and, thus, remains sound. The set of these conditions is called CSSM (Conditions for Sound Security Metrics) and is based on the verification of three properties of the evaluated metric:

- Well-defined, i.e., it measures components of the security level and is meaningful, objective, unbiased, and complete, to not miss any aspects of the definition needed to be effective as well as not to be too expensive to evaluate;
- Progressive, i.e., the metric expresses a value or set of values that coherently evolve together with the actual level of security so that progress towards an acceptable value of the metric is an indicator of improved security level;
- Strongly or weakly reproducible, i.e., it can be used in different environments and still produce comparable results.

The threat landscape characterizing cyber–physical systems (CPSs) is peculiar. In the literature, some recent studies have shown a growing interest in adopting metrics to measure security, showing that to choose what properties should be measured, the convergence of physical and cyber components must be taken into account: cyber threats can have tangible, real-world consequences, and the interconnectedness, also involving public infrastructure, amplifies the attack surface and the potential impact of attacks.

One model to summarize threats in CPS is STRIDE [\[24\]](#page-15-10): spoofing identity, tampering with data, repudiation of origin, information disclosure, denial of service, and elevation of privilege. Key aspects to address include understanding the possible consequences of attacks, determining the peculiar properties of CPS and the consequences that differ from traditional systems, and finding and testing security mechanisms applicable to CPS.

Still, the application of such models to define metrics in CPS is in its early stages. For example, in the work by Aigner et al. [\[25\]](#page-15-11), a thorough suitability evaluation of whether the selected metrics meet the conditions posed by CPS was conducted. The results indicate that while the metrics cover nearly all desired features, none of them fully address the entirety of the challenges proposed. The main concern stems from the fact that the analyzed metrics primarily focus on the specific elements of the system without adequately considering the emerging properties that arise from their composition, such as dependencies and side effects within system of systems contexts. Consequently, the emphasis remains primarily on vulnerabilities and attacks, overlooking important aspects of the CPS's overall security posture.

Other than context constraints, some security metrics were proven to be inadequate for their intended objectives, as shown in Yee et al. [\[26\]](#page-15-12). They either evaluate incorrect factors or fail to incorporate a sufficient number of relevant factors, consequently resulting in baseless and irrational conclusions. These problems can also arise from subjective rather than objective assessments, inaccurate estimates that cannot be reproduced, and distortions of actual measurements that lead to erroneous conclusions. To avoid those problems we need an adequate strategy for designing and testing security metrics.

3. Selection Methodology

As a result of our analysis of related works, we claim that effective metrics for CPS are needed but not available. In this section, we outline our proposed methodology for the systematic collection, selection, and validation of security metrics tailored for the specific subdomain comprising industrial cyber–physical systems. By restricting the domain, we can achieve the goal of selecting solid metrics, yet we strive to keep the process general enough to be extended to any CPS. We explain the approach with an algorithmic-style description (Algorithm [1\)](#page-5-0) highlighting the steps that lead to the final set of security metrics ready to be used.

The underlying logic is also depicted in Figure [1,](#page-4-0) which provides a summary of the steps of the algorithm, individually detailed hereinafter.

Figure 1. A graphical summarizationof our objective with the several steps required for the achievement of the set of security metrics for a specific domain.

Algorithm 1 An algorithmic view on the procedure to obtain the filtered and validated set of security metrics.

The first step of our methodology targets an extensive review of the existing literature to collect a comprehensive set of security metrics. We aim to encompass a wide spectrum of metrics, considering both quantitative and qualitative indicators, similar to the approaches presented in Section [2.](#page-1-0) This initial phase serves as the foundation for building a diverse pool of potential metrics.

Next, the metrics undergo a meticulous classification; from the literature analysis, we derive a set of relevant properties that are crucial for categorizing security metrics. These properties are discerned by pinpointing the relevant requirements in the context of (I)CPS security. The goal is to extract key features that will be the foundation for a nuanced classification.

We structure our classification with attributes that are drawn from previous works, mainly Villarrubia et al. [\[27\]](#page-15-13) and Savola et al. [\[28\]](#page-15-14), that propose taxonomies for security metrics. The selected attributes are:

- A name, that represents in a few words the meaning of the metric.
- A definition, that describes the metric and what it measures.
- A meaning, that summarizes the objective measurement of the metric and why it is useful.
- A weakness, that shows requirements, possible problems, or critical issues related to the metric, e.g., needing external support to calculate the metrics.
- A scope, that represents the field in which the metric is focused, e.g., network, device, user, organizations, system, etc.
- A result type, that can be quantitative if the metric gives a result in a numerical form or qualitative if the result is in a descriptive and discrete form (e.g., bad, normal, good).
- An automation field, that divides metrics into automatic, where the computation can be performed in an automatic way, or manual, where humans are required.
- A measurement field, that could be dynamic if the metric changes at runtime or *static* if the metric only changes with a new configuration.
- A construction, that can be modeled if the metrics need a model to be computed, i.e., an attack graph [\[29\]](#page-15-15), or measured if it represents a simple calculation that can be directly executed without models.

To restrict the outcome of the selection process, it is essential to conduct a detailed analysis of the application domain for each metric. In this step, we define a set of properties that describe metric usability and their relevance according to the intricacies of the domain.

To achieve reliability, effectiveness, and real-world applicability, the selected metrics then undergo validation using the CSSM. This robust validation framework ensures that the chosen metrics not only meet theoretical expectations but also demonstrate practical utility and reproducibility while ensuring convergence of values when we reach a supposedly secure condition. In summary, we claim that our proposed selection approach can output useful and relevant metrics by combining width, from the extensive literature analysis, and depth, from the tailored definition of criteria and robust validation of the results.

4. Application: Metrics Selection for ICPSs

To gather a comprehensive collection of security metrics specifically tailored for ICPSs, we applied the proposed methodology within this specific context. Our collaboration with domain experts proved invaluable, as their expertise facilitated the meticulous maintenance of accurate and consistent selections throughout the process. We made available publicly [\(https://doi.org/10.5281/zenodo.10142113,](https://doi.org/10.5281/zenodo.10142113) accessed on 19 March 2024) the initial dataset, the final metric set, and all the reduction steps.

4.1. Metrics Classification

In the cited work by Longueira-Romero et al. [\[6\]](#page-14-5), before filtering the metrics for a set applicable only in the embedded systems field, the authors followed a variation of the search and selection strategy presented in Section [2:](#page-1-0)

- Initially, they chose the data sources for the metrics, resulting in conference proceedings and academic journals from IEEE Xplore, Elsevier, AMC Digital Library, Springer, and Google Scholar search engine.
- Then, they gathered the metrics and labeled them with definition, scale, scope, automation, and measurement attributes.
- Finally, they filtered the metrics.

We decided to start the basis of our set of metrics with the dataset presented in their work before the reductions that they specifically applied for ES, which allowed us to initialize our set with more than 500 metrics taken from different sources in the literature [\[10,](#page-14-9)[30](#page-15-16)[–32\]](#page-15-17). We then extended their discovery work by further exploring the literature and integrating the dataset. We searched over online databases such as IEEE Xplore, Elsevier, AMC Digital Library, and Springer, as well as conference proceedings and academic journals, including Google Scholar search engine, to gather security metrics. Our search terms encompassed keywords like "security metric", "icps" and "assessment" in addition to relevant synonyms. Our inclusion criteria prioritized papers focusing on security measurements or metrics, with a preference for surveys on collecting security metrics, and those primarily concerned with measuring security. Using tools such as the search engine Scholar with 14 more metrics from Boyer et al. [\[33\]](#page-15-18) and 29 from Bhol et al. [\[34\]](#page-15-19). After pruning repeated entries or not-referenced ones, we counted a total of 278 security metrics.

We decided to start the basis of our set of metrics with the dataset presented in their work before the reductions that they specifically applied for ES, which allowed us to initialize our set with more than 500 metrics taken from different sources in the literature [\[10,](#page-14-9)[30](#page-15-16)[–32\]](#page-15-17). We then extended their discovery work by further exploring the literature and integrating the dataset. Our search terms encompassed keywords like "security metric", "security assessment", "icps" and "cps" in addition to relevant synonyms searched over online databases such as IEEE Xplore, Elsevier, AMC Digital Library, and Springer, as well as conference proceedings and academic journals, including Google Scholar search engine. Our inclusion criteria prioritized papers containing security metrics with a definition sufficiently clear to be applied in real use-case scenarios. We finally gather 14 metrics from Boyer et al. [\[33\]](#page-15-18) and 29 from Bhol et al. [\[34\]](#page-15-19). After pruning repeated entries or not-referenced ones, we counted a total of 278 security metrics.

4.2. Domain Analysis

We structured our dataset by classifying metrics with the most common attributes present in the literature, shown in Section [2.](#page-1-0) Classifying metrics according to certain characteristics enables us to choose them according to criteria that depend on the context. In our case, we refer to the domain of ICPSs, so it is necessary to study their intrinsic characteristics before proposing a set of metrics that fully capture the security issues that may arise.

ICPSs are composed of interconnected cyber and physical components that monitor and manage physical processes. They are responsible for the safety and operations of the industrial process, which implies the management of heterogeneous hardware and software. They include devices such as sensors, actuators, Supervisory Control And Data Acquisition (SCADA) systems, human–machine interfaces (HMIs), and dedicated subsystems such as programmable logic controllers (PLCs) [\[35\]](#page-15-20). This heterogeneity obviously translates into system complexity, which implies more effort to manage and prevent anomalies. In addition, ICPS networks employ a wide range of protocols, depending on the specific objectives of each system. Real-time constraints and legacy hardware are two of the most important challenges that industrial protocols are specifically made to address. The Purdue Enterprise Reference Architecture [\[36\]](#page-15-21) is the reference networking architecture for ICPS systems, adopted in the ANSI/ISA-95 standard [\[37\]](#page-15-22), and divides ICPS networks into three logical segments: the lower layer is the manufacturing zone, also known as operational technology (OT), while the upper constitute the enterprise zone, also referred to as information technology (IT), with a demilitarized zone of convergence between them. The OT network includes hardware and software used to monitor and manage industrial equipment, assets, processes, and events. On the other side, the traditional information technology (IT) network contains workstations, databases, and other typical machines used to manipulate information. From this perspective, IT systems' main concerns are about the confidentiality and integrity of the data, while for the OT part, availability is instead fundamental since it can guarantee human safety and fault tolerance [\[38\]](#page-15-23).

4.3. Metrics Filtering

Starting from the domain analysis of ICPSs and the attributes of each metric, we established the following inclusion criteria for metrics filtering:

- The definition of the metric must be applicable to IT or OT networks, components, protocols, and devices.
- The meaning attribute of the metric must explicitly declare an objective measurement related at least to one of the security properties of confidentiality, integrity, or availability.
- The weakness attribute of the metrics must be related to problems, issues, or requirements that can be resolved inside the ICPS domain.
- The scope attribute must be of the type "network", "device", "system", or "user".

To match the aforementioned inclusion criteria, we performed two reduction steps of the original dataset. In the first reduction, we filtered only metrics that have as scope "network", "device", "system", or "user". Then, for every metric within those scopes, we carefully checked the other three inclusion criteria, resulting in a second reduction.

To provide a demonstration of the second reduction, we consider the two metrics shown in Table [1.](#page-8-0) The first one is the "infection rate" [\[39\]](#page-15-24): this metric is denoted by the average number of hosts that can be infected per unit of time by one infected host during the early stage of worm propagation. We consider this metric applicable to our case study because measuring the spreading of malware is crucial [\[40\]](#page-15-25) in evaluating the cybersecurity level of an industrial system, in particular for the IT network. On the other hand, the "ISP badness metric" [\[41\]](#page-15-26) does not match our target, because its definition (as reported in Table [1\)](#page-8-0) does not regard to any extent IT or OT components.

Metric Definition Ref. Match IC Infection Rate Average number of computers that can be infected by a compromised computer (per time unit) at the early stage of spreading [\[39\]](#page-15-24) Yes ISP badness metric Quantifies the effect of spam from one ISP or autonomous system on the rest of the Internet comparing the "spamcount" with its "disconnectability" [\[41\]](#page-15-26) No

Table 1. Example metrics that match or do not match the first three inclusion criteria (IC) regarding the second reduction step.

As the final objective of our project is to have a set of metrics that are actually usable in real experiments, we decided to not limit our filtering merely to the selection of metrics that are domain-relevant but to go beyond that and find only metrics that are valid and feasible to use in ICPSs. For this reason, we applied the CSSM framework to every metric that we filtered in the previous reduction, checking if they respect all the three requirements (well-defined, progressive, and reproducible), explaining the verification process attended and the reasons that led to the specific outcome. Yee et al. [\[23\]](#page-15-9) suggest evaluating the completeness condition of a metric in a team where everyone has the "big picture" in mind and can share and compare opinions; thus, we evaluated the metrics individually in the research group and then compared ideas collectively.

To explain this step, let us dive into four validation examples, as shown in Table [2:](#page-9-0) even though they are all usable in our use-case scenario, three of them do not respect one requirement of the CSSM and only one passes all the tests. The first one is "vulnerability lifetime", i.e., the amount of time that a vulnerability remains in the system. This metric does not pass the well-defined (WD) condition because it is not always possible to find the exact moment the vulnerability enters the system. Meneely et al. [\[42\]](#page-16-0) argue that what is referred to in the literature as vulnerability-contributing commits (VCCs), which are repository commits that lead to the introduction of a vulnerability after release, do not always guarantee to rightly reflect the security posture of the analyzed system, as demonstrated in Alexopoulos et al. [\[43\]](#page-16-1). The second one is the "network maliciousness metric", which measures the amount of blacklisted IP addresses in a network. This metric does not pass the progressive (P) condition since the amount of blacklisted IPs does not necessarily indicate the security level of the system. The third one is "worst-case loss", which measures the maximum dollar value of the loss that could happen in the system. This metric does not pass the reproducible (R) condition because the way to assess this amount is just an estimation based on various factors that are not deterministic, and thus, not reproducible in different systems with comparable results. The last one is "VEAbility". This metric is based on the aggregation of the CVSS scores on a certain network configuration and passes all the conditions: it is a complete measure of the security of the system, it is a score that converges to a secure level and it is based on a standardized and reproducible assessment.

Table 2. Example metrics that respect (V) or not (X) the three CSSM conditions: well-defined (WD), progressive (P), and reproducible (R).

Lastly, the final reduction involved a confirmation that there were no overlapping metrics among the set of metrics. To ensure this, we considered equivalent metrics that share the same objective and concept, other than the same attributes: scope, result types, automation, measurement, construction, and type. Then, we merged equivalent metrics to obtain the final set: the resulting metrics prove to be valid, as the CSSM still applies, thanks to the rule demonstrated in Yee et al. [\[23\]](#page-15-9) that states that "an additive aggregate security metric is sound if all of the security metrics that are sums are sound".

4.4. Results

In this section, we present the results of our work: the security metrics that have successfully gone through all stages of our meticulous selection and validation process. We outline the characteristics and properties of the validated metrics, providing a foundation for their practical application in the realm of ICPS.

Figure [2](#page-10-0) delineates the distribution of security metrics across distinct categories, namely, vulnerability, attack, defense, and situation, at the beginning and after the stage of filtering and validation. This visualization offers insights into the evolution of metric

categories throughout the process of selection, highlighting the adaptability and relevance of each category as metrics progress through successive steps.

We begin our selection process with a total of 278 security metrics, deleting entires that are not useful or not valid after each step; the results are shown in Figure [3.](#page-11-0) As we can observe, the 3-step CSSM validation prunes the majority of not-valid entries with the first well-defined condition check: this can be explained by the fact that several metrics that are found not to satisfy the progressive and reproducible conditions are also not well-defined in the first place.

Distribution of metrics for every reduction

Figure 2. Percentage of security metrics of different categories after every step of our approach of filtering and validation.

After the validation, we collected a total of 32 security metrics that represented our final set. To understand the quality of evaluation of our metrics combined together, we evaluated which aspects of confidentiality, integrity, and availability (CIA) were taken into account: 87.5% of our metrics covered all the three aspects of CIA, whereas the remaining 12.5% covered only one aspect. Table [3](#page-12-0) shows this result: a list of all the metrics that we found, with the aspects of CIA that they cover. These metrics demonstrated resilience and applicability across the specified criteria, positioning them as reliable indicators of security within the context of our research.

A final representation of our results is shown in Figure [4:](#page-11-1) in the initial set, the partitioning between static and dynamic metrics was not balanced; however, as we performed each step of the filtering and validation process, the difference in quantities between static and dynamic security metrics improved, reaching an almost balanced partitioning.

Amount of metric after every step

Figure 3. Number of security metrics after every step of our approach of filtering and validation with CSSM (well-defined WD, progressive P, reproducible R).

Distribution of Dynamic and Static metric after every reduction

Figure 4. Percentage of dynamic and static security metrics after every step of our approach of filtering and validation.

Table 3. The list of the final set of security metrics that we obtained as a result of the collection, filtering, and validation for the ICPS domain. The "X" in the CIA columns indicates that the relative property is considered.

5. Discussion and Limitations

While our endeavor aimed to obtain a comprehensive set of security metrics for ICPSs, which is an objective both ambitious and complicated, certain limitations and observations merit consideration. Our systematic methodology heavily relied on the available literature for metric collection. The completeness of our set is contingent upon the extent and depth of existing publications in the field. Moreover, the metrics identified may not comprehensively capture emerging threats or changes in the ICPS landscape over time. This is due to the ICPS domain that is inherently dynamic, evolving with technological advancements and consisting of complex interconnection of systems, each one with their own vulnerabilities and diverse technological constraints.

The process of classifying security metrics involves interpreting and applying criteria to categorize them into distinct types. Despite having established specific criteria and performed an analysis of the definition for each one of them, this operation can still be influenced by subjective judgments. Different individuals may interpret the criteria differently, leading to potential variations in how metrics are categorized: it is important to acknowledge that the subjectivity in classification does not undermine the validity of the overall methodology. Instead, it underscores the need for transparency in the classification process and a recognition of the interpretative aspects involved. Future refinement of classification criteria and potentially leveraging more consensus-building approaches among experts in the field could contribute to minimizing subjectivity and enhancing the reproducibility of the classification process.

Moreover, the coverage of the resulting metrics remains confined to specific technical aspects. The deployment and operation of ICPSs, and of CPSs in general, nonetheless, extend beyond narrow technical considerations, encompassing a broader range of concerns. This includes social dimensions, safeguarding fundamental human rights like privacy and ethical considerations, ensuring physical safety, and investigating interactions with the broader landscape of threats while integrating intelligence about them. An interesting conclusion is that in our set only 3% fall in the user category: this evidence highlights how little importance is given to the identification of user-related issues in security evaluations, although the user is usually seen as the weakest link in cybersecurity [\[51\]](#page-16-10).

6. Conclusions and Future Work

In the pursuit of establishing a comprehensive suite of security metrics tailored for ICPSs and ensuring a nuanced understanding of its unique challenges, this study gives a twofold contribution to the field.

On the one hand, we define a systematic multi-stage methodology proceeding from the broadest collection of metrics through a validated classification, selection, and filtering process, to ensure the reliability, reproducibility, and practical applicability of the selected metrics. On the other hand, by testing such a methodology, we obtain a substantial compilation of ICPS-specific security metrics, reflecting a broad spectrum of considerations.

Such contributions serve as a foundational resource for further research, aiding practitioners, researchers, and decision-makers in enhancing the security posture of ICPSs.

We adopted CSSM as a robust validation framework, yet its application may not cover all possible dimensions of security. Future research could explore additional validation mechanisms to enhance the thoroughness of the validation process. The threat landscape is ever-evolving; hence, ongoing efforts are necessary to adapt and expand the set of metrics to address emerging security challenges. Future works could also use the proposed set of metrics to perform a long-term testing phase inside companies to empirically evaluate the utility of the metrics that we have found.

Author Contributions: Conceptualization, G.G., L.R., A.M., and A.A.S.; methodology, G.G., L.R., and A.A.S.; validation G.G., L.R., and A.M.; writing—original draft, G.G., L.R., and A.M.; writing—review and editing, A.M., A.A.S., F.C., and M.P.; supervision, A.M., F.C., and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by project SERICS (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union–NextGenerationEU.

Data Availability Statement: The data presented in this study are openly available in Zenodo at [https://doi.org/10.5281/zenodo.10142113,](https://doi.org/10.5281/zenodo.10142113) accessed on 19 March 2024.

Acknowledgments: This work was partially supported by the IGNITE5.0 project funded by the Emilia-Romagna region in the framework of the POR-FESR 2021–2027 program (the EU program for regional development); Giacomo Gori was supported by the National PhD in Cybersecurity (IMT Lucca); Amir Al Sadi was supported by the National Operating Program Research and Innovation (PON 2014–2020), action IV.5 PhD on Green subjects, by the Advanced Studies Institution (ISA) and by the International PhD College.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

References

- 1. Azuwa, M.; Ahmad, R.; Sahib, S.; Shamsuddin, S. Technical security metrics model in compliance with ISO/IEC 27001 standard. *Int. J. Cyber-Secur. Digit. Forensics* **2012**, *1*, 280–288.
- 2. Chew, E.; Swanson, M.; Stine, K.M.; Bartol, N.; Brown, A.; Robinson, W. *Sp 800–55 Rev. 1. Performance Measurement Guide for Information Security*; NIST: Gaithersburg, MD, USA, 2008.
- 3. Tran, J.L. Navigating the Cybersecurity Act of 2015. *Chap. L. Rev.* **2016**, *19*, 483.
- 4. Philippou, E.; Frey, S.; Rashid, A. Contextualising and aligning security metrics and business objectives: A GQM-based methodology. *Comput. Secur.* **2020**, *88*, 101634. [\[CrossRef\]](http://doi.org/10.1016/j.cose.2019.101634)
- 5. Wang, L.; Jajodia, S.; Singhal, A. *Network Security Metrics*; Springer: Berlin/Heidelberg, Germany, 2017.
- 6. Longueira-Romerc, Á.; Iglesias, R.; Gonzalez, D.; Garitano, I. How to quantify the security level of embedded systems? A taxonomy of security metrics. In Proceedings of the 2020 IEEE 18th International Conference on Industrial Informatics (INDIN), Warwick, UK, 20–23 July 2020; IEEE: New York, NY, USA, 2020; Volume 1, pp. 153–158.
- 7. Doran, G.T. There's a SMART way to write management's goals and objectives. *Manag. Rev.* **1981**, *70*, 35–36.
- 8. Brotby, W.K.; Hinson, G. *Pragmatic Security Metrics: Applying Metametrics to Information Security*; CRC Press: Boca Raton, FL, USA, 2013.
- 9. Savola, R.M. Quality of security metrics and measurements. *Comput. Secur.* **2013**, *37*, 78–90. [\[CrossRef\]](http://dx.doi.org/10.1016/j.cose.2013.05.002)
- 10. Pendleton, M.; Garcia-Lebron, R.; Cho, J.H.; Xu, S. A survey on systems security metrics. *ACM Comput. Surv. (CSUR)* **2016**, *49*, 1–35. [\[CrossRef\]](http://dx.doi.org/10.1145/3005714)
- 11. Hauet, J.P. ISA99/IEC 62443: A solution to cyber-security issues? In Proceedings of the ISA Automation Conference, Jeju Island, Republic of Korea, 17–21 October 2012.
- 12. Sultan, K.; En-Nouaary, A.; Hamou-Lhadj, A. Catalog of metrics for assessing security risks of software throughout the software development life cycle. In Proceedings of the 2008 International Conference on Information Security and Assurance (ISA 2008), Busan, Republic of Korea, 24–26 April 2008; IEEE: New York, NY, USA, 2008; pp. 461–465.
- 13. Morrison, P.; Moye, D.; Pandita, R.; Williams, L. Mapping the field of software life cycle security metrics. *Inf. Softw. Technol.* **2018**, *102*, 146–159. [\[CrossRef\]](http://dx.doi.org/10.1016/j.infsof.2018.05.011)
- 14. Ani, U.P.D.; He, H.; Tiwari, A. A framework for Operational Security Metrics Development for industrial control environment. *J. Cyber Secur. Technol.* **2018**, *2*, 201–237. [\[CrossRef\]](http://dx.doi.org/10.1080/23742917.2018.1554986)
- 15. Fernández-Alemán, J.L.; Señor, I.C.; Lozoya, P.Á.O.; Toval, A. Security and privacy in electronic health records: A systematic literature review. *J. Biomed. Inform.* **2013**, *46*, 541–562. [\[CrossRef\]](http://dx.doi.org/10.1016/j.jbi.2012.12.003)
- 16. Sundararajan, A.; Sarwat, A.I.; Pons, A. A survey on modality characteristics, performance evaluation metrics, and security for traditional and wearable biometric systems. *ACM Comput. Surv. (CSUR)* **2019**, *52*, 39. [\[CrossRef\]](http://dx.doi.org/10.1145/3309550)
- 17. Hong, J.B.; Enoch, S.Y.; Kim, D.S.; Nhlabatsi, A.; Fetais, N.; Khan, K.M. Dynamic security metrics for measuring the effectiveness of moving target defense techniques. *Comput. Secur.* **2018**, *79*, 33–52. [\[CrossRef\]](http://dx.doi.org/10.1016/j.cose.2018.08.003)
- 18. Manadhata, P.K.; Wing, J.M. An attack surface metric. *IEEE Trans. Softw. Eng.* **2010**, *37*, 371–386. [\[CrossRef\]](http://dx.doi.org/10.1109/TSE.2010.60)
- 19. Roundy, K.A.; Miller, B.P. Binary-code obfuscations in prevalent packer tools. *ACM Comput. Surv. (CSUR)* **2013**, *46*, 1–32. [\[CrossRef\]](http://dx.doi.org/10.1145/2522968.2522972)
- 20. Zhan, Z.; Xu, M.; Xu, S. A characterization of cybersecurity posture from network telescope data. In Proceedings of the Trusted Systems: 6th International Conference, INTRUST 2014, Beijing, China, 16–17 December 2014; Revised Selected Papers 6; Springer: Berlin/Heidelberg, Germany, 2015; pp. 105–126.
- 21. Zhang, J.; Durumeric, Z.; Bailey, M.; Liu, M.; Karir, M. On the Mismanagement and Maliciousness of Networks. In Proceedings of the NDSS, San Diego, CA, USA, 23–26 February 2014; Volume 14, pp. 23–26.
- 22. Ahmed, R.K.A. Security metrics and the risks: An overview. *Int. J. Comput. Trends Technol.* **2016**, *41*, 106–112. [\[CrossRef\]](http://dx.doi.org/10.14445/22312803/IJCTT-V41P119)
- 23. Yee, G.O. Improving the Derivation of Sound Security Metrics. In Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), Los Alamitos, CA, USA, 27 June–1 July 2022; IEEE: New York, NY, USA, 2022; pp. 1804–1809.
- 24. Shostack, A. *Threat Modeling: Designing for Security*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
- 25. Aigner, A.; Khelil, A. A Benchmark of Security Metrics in Cyber-Physical Systems. In Proceedings of the 2020 IEEE International Conference on Sensing, Communication and Networking (SECON Workshops), Como, Italy, 22–26 June 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.
- 26. Yee, G.O. Designing sound security metrics. *Int. J. Syst. Softw. Secur. Prot.* **2019**, *10*, 1–21. [\[CrossRef\]](http://dx.doi.org/10.4018/IJSSSP.2019010101)
- 27. Villarrubia, C.; Fernández-Medina, E.; Piattini, M. Towards a Classification of Security Metrics. In Proceedings of the WOSIS, Porto, Portugal, 1 April 2004; pp. 342–350.
- 28. Savola, R. Towards a security metrics taxonomy for the information and communication technology industry. In Proceedings of the International Conference on Software Engineering Advances (ICSEA 2007), Cap Esterel, France, 25–31 August 2007; IEEE: New York, NY, USA, 2007; p. 60.
- 29. Ou, X.; Boyer, W.F.; McQueen, M.A. A scalable approach to attack graph generation. In Proceedings of the 13th ACM Conference on Computer and Communications Security, Alexandria, VA, USA, 30 October–3 November 2006; pp. 336–345.
- 30. Ramos, A.; Lazar, M.; Holanda Filho, R.; Rodrigues, J.J. Model-based quantitative network security metrics: A survey. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 2704–2734. [\[CrossRef\]](http://dx.doi.org/10.1109/COMST.2017.2745505)
- 31. Enoch, S.Y.; Hong, J.B.; Ge, M.; Kim, D.S. Composite metrics for network security analysis. *arXiv* **2020**, arXiv:2007.03486.
- 32. Morrison, P.; Moye, D.; Williams, L.A. *Mapping the Field of Software Security Metrics*; Technical Report; Department of Computer Science, North Carolina State University: Raleigh, NC, USA, 2014.
- 33. Boyer, W.; McQueen, M. Ideal based cyber security technical metrics for control systems. In Proceedings of the Critical Information Infrastructures Security: Second International Workshop, CRITIS 2007, Málaga, Spain, 3–5 October 2007; Revised Papers 2; Springer: Berlin/Heidelberg, Germany, 2008; pp. 246–260.
- 34. Bhol, S.G.; Mohanty, J.; Pattnaik, P.K. Taxonomy of cyber security metrics to measure strength of cyber security. *Mater. Today Proc.* **2023**, *80*, 2274–2279. [\[CrossRef\]](http://dx.doi.org/10.1016/j.matpr.2021.06.228)
- 35. Conti, M.; Donadel, D.; Turrin, F. A survey on industrial control system testbeds and datasets for security research. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 2248–2294. [\[CrossRef\]](http://dx.doi.org/10.1109/COMST.2021.3094360)
- 36. Williams, T.J. The Purdue enterprise reference architecture. *Comput. Ind.* **1994**, *24*, 141–158. [\[CrossRef\]](http://dx.doi.org/10.1016/0166-3615(94)90017-5)
- 37. ISA95, Enterprise-Control System Integration. Available online: [https://www.isa.org/standards-and-publications/isa](https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95)[standards/isa-standards-committees/isa95](https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa95) (accessed on 19 March 2024).
- 38. Zhang, K.; Shi, Y.; Karnouskos, S.; Sauter, T.; Fang, H.; Colombo, A.W. Advancements in industrial cyber-physical systems: An overview and perspectives. *IEEE Trans. Ind. Inform.* **2022**, *19*, 716–729. [\[CrossRef\]](http://dx.doi.org/10.1109/TII.2022.3199481)
- 39. Chen, Z.; Ji, C. Measuring network-aware worm spreading ability. In Proceedings of the IEEE INFOCOM 2007—26th IEEE International Conference on Computer Communications, Barcelona, Spain, 6–12 May 2007; IEEE: New York, NY, USA, 2007; pp. 116–124.
- 40. Goebel, J.; Holz, T.; Willems, C. Measurement and analysis of autonomous spreading malware in a university environment. In Proceedings of the Detection of Intrusions and Malware, and Vulnerability Assessment: 4th International Conference, DIMVA 2007, Lucerne, Switzerland, 12–13 July 2007; Proceedings 4; Springer: Berlin/Heidelberg, Germany, 2007; pp. 109–128.
- 41. Johnson, B.; Chuang, J.; Grossklags, J.; Christin, N. Metrics for Measuring ISP Badness: The Case of Spam: (Short Paper). In Proceedings of the Financial Cryptography and Data Security: 16th International Conference, FC 2012, Kralendijk, Bonaire, 27 Februray–2 March 2012; Revised Selected Papers 16; Springer: Berlin/Heidelberg, Germany, 2012; pp. 89–97.
- 42. Meneely, A.; Srinivasan, H.; Musa, A.; Tejeda, A.R.; Mokary, M.; Spates, B. When a patch goes bad: Exploring the properties of vulnerability-contributing commits. In Proceedings of the 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, Baltimore, MD, USA, 10–11 October 2013; IEEE: New York, NY, USA, 2013; pp. 65–74.
- 43. Alexopoulos, N. New Approaches to Software Security Metrics and Measurements. Ph.D. Thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2022.
- 44. Tupper, M.; Zincir-Heywood, A.N. VEA-bility security metric: A network security analysis tool. In Proceedings of the 2008 Third International Conference on Availability, Reliability and Security, Barcelona, Spain, 4–7 March 2008; IEEE: New York, NY, USA, 2008; pp. 950–957.
- 45. Lanotte, R.; Merro, M.; Munteanu, A.; Tini, S. Formal impact metrics for cyber-physical attacks. In Proceedings of the 2021 IEEE 34th Computer Security Foundations Symposium (CSF), Dubrovnik, Croatia, 21–25 June 2021; IEEE: New York, NY, USA, 2021; pp. 1–16.
- 46. Zhang, M.; Wang, L.; Jajodia, S.; Singhal, A.; Albanese, M. Network diversity: A security metric for evaluating the resilience of networks against zero-day attacks. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 1071–1086. [\[CrossRef\]](http://dx.doi.org/10.1109/TIFS.2016.2516916)
- 47. Gu, G.; Fogla, P.; Dagon, D.; Lee, W.; Skorić, B. Measuring intrusion detection capability: An information-theoretic approach. In Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security, Taipei, Taiwan, 21–24 March 2006; pp. 90–101.
- 48. Wang, L.; Jajodia, S.; Singhal, A.; Noel, S. k-zero day safety: Measuring the security risk of networks against unknown attacks. In Proceedings of the Computer Security—ESORICS 2010: 15th European Symposium on Research in Computer Security, Athens, Greece, 20–22 September 2010; Proceedings 15; Springer: Berlin/Heidelberg, Germany, 2010; pp. 573–587.
- 49. Sharma, D.P.; Enoch, S.Y.; Cho, J.H.; Moore, T.J.; Nelson, F.F.; Lim, H.; Kim, D.S. Dynamic security metrics for software-defined network-based moving target defense. *J. Netw. Comput. Appl.* **2020**, *170*, 102805. [\[CrossRef\]](http://dx.doi.org/10.1016/j.jnca.2020.102805)
- 50. Demme, J.; Martin, R.; Waksman, A.; Sethumadhavan, S. Side-channel vulnerability factor: A metric for measuring information leakage. In Proceedings of the 2012 39th Annual International Symposium on Computer Architecture (ISCA), Portland, OR, USA, 9–13 June 2012; Volume 40, pp. 106–117.
- 51. Jalkanen, J. Is Human the Weakest Link in Information Security?: Systematic Literature Review. Master's Thesis, University of Jyväskylä, Jyväskylän yliopisto, Finland, 2019.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.