Risk Assessments Considering Safety, Security, and Their Interdependencies in OT Environments

Siegfried Hollerer
siegfried.hollerer@tuwien.ac.at
Institute of Computer Engineering, TU Wien
Vienna, Austria

Thilo Sauter
thilo.sauter@tuwien.ac.at
Inst. of Computer Techn., TU Wien
Dept. for Integrated Sensor Systems, Danube Univ. Krems

Wolfgang Kastner
wolfgang.kastner@tuwien.ac.at
Institute of Computer Engineering, TU Wien
Vienna, Austria

ABSTRACT

Information Technology (IT) and Operational Technology (OT) are converging further, which increases the number of interdependencies of safety and security risks arising in industrial architectures. Cyber attacks interfering safety functionality may lead to serious injuries as a consequence. Intentionally triggering a safety function may introduce a security vulnerability during the emergency procedure, e.g., by opening emergency exit doors leading to enabling unauthorized physical access. This paper introduces a risk evaluation methodology to prioritize and manage identified threats considering security, safety, and their interdepedencies. The presented methodology uses metrics commonly used in the industry to increase its applicability and enable the combination with other risk assessment approaches. These metrics are Common Vulnerability Scoring System (CVSS), Security Level (SL) from the standard IEC 62443 and Safety Integrity Level (SIL) from the standard IEC 61508. Conceptional similarities of those metrics are considered during the risk calculation, including an identified relation between CVSS and SL. Besides this relation, the skill level and resources of threat actors, threats enabling multiple identified attacks, the SIL of safety-relevant components affected, business criticality of the targeted asset, and the SL-T of the zone targeted by the attack are considered for risk evaluation. The industrial architecture to be analyzed is separated into zones and conduits according to IEC 62443, enabling the analyzed system to be compliant with its requirements.

CCS CONCEPTS

• **Hardware** → *Safety critical systems*; • **Security and privacy** → **Distributed systems security**.

KEYWORDS

Threat Modeling, OT Security, Safety, IT / OT convergence

ACM Reference Format:
Siegfried Hollerer, Thilo Sauter, and Wolfgang Kastner. 2022. Risk Assessments Considering Safety, Security, and Their Interdependencies in OT Environments. In *The 17th International Conference on Availability, Reliability and Security (ARES 2022), August 23–26, 2022, Vienna, Austria.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3538969.3543814

ARES 2022, August 23–26, 2022, Vienna, Austria

ACM ISBN 978-1-4503-9670-7/22/08.
https://doi.org/10.1145/3538969.3543814

1 INTRODUCTION

The domains of Information Technology (IT) and Operational Technology (OT) are continuing to converge in industrial automation. They used to be isolated from each other, as the Purdue Enterprise Reference Architecture (PERA) [14, 22] schematically suggests. IT is located on the top level of this reference architecture containing enterprise systems, office networks, and software to process and distribute data of production plants. All other levels are considered as OT containing hardware, industrial communication systems, and software components to control and observe the technical processes of machines. The convergence of IT and OT leads to challenges in addressing both protection goals of industrial architectures, which are security and safety.

Originating from the IT domain, security aims to protect the confidentiality, integrity, and availability (also known as "CIA-triad") of data to protect against cyber attacks. Coming from the OT domain, safety aims to ensure functional safety and resilience to protect humans and the environment from an undesirable operation that may lead to injuries and physical damage. Safety and security are interdependent [4]. For example, the manipulation of a Safety Instrumented Systems (SIS), a device meant to increase safety, may lead to not being able to execute the safety function when needed [6, 15]. This interference with the safety function of the SIS may injure people or damage the industrial architecture as a consequence. The other way around, the intentional execution of a safety function may be used to introduce a security vulnerability based on their interdependency [4]. For example, pushing an emergency stop button can open emergency doors during the emergency process, enabling unauthorized physical access of an attacker on-site [3, 9].

A methodology to identify, evaluate and treat risks arising from safety, security, and their interdependence is needed to prevent the occurrence of such events in a specific industrial architecture referred to as System under Consideration (SUC) [9]. The risk identification was elaborated in another work [4], therefore this paper focuses on the other phases of risk assessment (analysis and evaluation). The analysis is based on a state-of-the-art methodology to identify attacks [2] with the extension to consider threat actors for each attack. The proposed risk evaluation serves as the key contribution of this work. Metrics commonly used in the industry [4] are used, which eases the combination with other existing risk assessment approaches. The introduced calculation and evaluation of these metrics provide a system-tailored prioritization of attacks and their underlying threats to be treated considering safety requirements, security requirements, and their interdependencies in the defined SUC. Therefore, the introduced risk evaluation and treatment of the threat modeling approach is in accordance with

the standard IEC 62443 [11] while it does not suffer from the lack of a bi-directional assessment between security threats and their interdependence with safety functions that may occur when using approaches focusing on security only in OT environments (cf. e.g., [17, 20]).

This paper has the following structure: Section 2 provides background information about attack modeling and commonly used metrics for evaluating security and safety risks, namely Common Vulnerability Scoring System Version 3.1 (CVSS), Security Level (SL), and Safety Integrity Level (SIL). Section 3 introduces a risk evaluation scheme that considers attacks based on identified threats, threat actors, the SIL of safety-relevant OT components, the business-criticality of OT components, and the predefined Security Level Target (SL-T) of the corresponding zone or conduit. Section 4 describes a use case created based on the results of a stakeholder analysis [8]. Section 5 applies the introduced methodology to the use case, while Section 6 discusses the obtained results. Finally, Section 7 concludes this work.

2 BACKGROUND AND PREVIOUS WORK

Taking an integrated view on safety and security is not specific to factory or process automation. It is relevant also in other domains, such as railway and train control [1, 23]. The risk analysis methodologies developed there are however domain-specific and difficult to transfer to the OT domain.

This work is based on previous publications [4, 9] that introduce and apply an OT threat modeling approach addressing safety, security, and their interdependencies in a fog computing environment. Figure 1 illustrates the phases of this threat modeling approach. The phase *risk identification* is already discussed in [4] while [9] provides a survey and evaluation of existing threat modelling approaches for OT environments. Therefore, this paper elaborates on the remaining phases *risk analysis*, *evaluation* and *risk treatment*.

The proposed approach needs human experts to define a SUC prior to the actual analysis. All other phases are designed to be done by a (semi-)automated software system using an ontology that defines trust boundaries, safety requirements, security requirements, and interdependencies between safety and security requirements.

2.1 Attack modelling

In the phase *attack modeling*, the identified threats and interdependencies from *risk identification* are linked to attacks and threat actors to enable prioritisation of these threats. The alignment of threats to attacks is based on the Mitre Att&ck framework [2]. Various types of attacks are considered using this framework, including initial access to the OT architecture (e.g., via drive-by compromise, phishing, or exploitation of remote services), execution of malicious code (e.g., using command-line interfaces, APIs, scripting, or user execution), and lateral movement (e.g., using default credentials). The used types of threat actors are based on [3] and [18].

- Basic user: This is the basic unstructured hacker, cracker, or hobbyist and someone who uses established and potentially also automated techniques to attack a system. The adversaries of this kind have average access to hardware, software, and Internet connectivity - purchasable with average personal funds or theft from their employers.
- Insider: The insiders are basic users with the difference in the employment position inside the company. The privileges they own tightly correlate to their employment position (user, administrator, supervisor).
- Hacktivist: The hacktivists use their hacking abilities to promote a political agenda. Their intentions are often related to freedom of information.
- Cybercriminal: This is the ”black hat” type of hacker, i.e., an attacker with high knowledge and skills but criminal intentions. This category of attackers exploit known vulnerabilities and may find zero-days on their own. Their goals include blackmailing, espionage, and sabotage.
- Nation-state attacker: This type of attacker is sponsored by a government. They possibly belong to a state organization for carrying out offensive cyber operations. Typical targets are general intelligence and public infrastructure systems, traffic management, and power or water systems.
- Terrorist: The terrorist or cyber-terrorist is a politically motivated attacker who uses computers to cause severe disruption or widespread fear.

2.2 Risk evaluation

The phase *risk evaluation* of the introduced threat modeling approach [9] uses a combination of multiple commonly used metrics to consider both security-relevant and safety-relevant risks. The metrics used are CVSS, SL, and SIL.

2.2.1 *Common Vulnerability Scoring System Version 3.1.*

CVSS score [5] is a metric commonly used in the industry for rating cyber security risks. Several compliance bodies recommend the usage of CVSS. This includes the U.S. government (e.g., National Institute of Standards and Technology (NIST) SP 800-115 and 800-43) and the global payment card industry (e.g., Payment Card Industry Data Security Standard (PCI-DSS) in the regulation on Approved Scanning Vendor (ASV)s). Despite the high acceptance in the industry, flaws were identified in CVSS [8, 19], such as:

- Insufficient addressing of technical and human-organizational context
- Potential material consequences against life or property (safety) of exploiting the vulnerability are not considered
- Operational scoring problems arise (e.g., inconsistencies due to design flaws and missing formalization of the scoring algorithm)

Since CVSS does not consider potential safety-relevant consequences, the CVSS-based Robot Vulnerability Scoring System (RVSS) includes safety as a metric influencing the impact of the vulnerability. The metric (year) of RVSS is introduced to evaluate the time since the vulnerability was first reported because security updates at OT components tend to be not installed for a longer time frame than in IT. Additionally, the metric attack vector in CVSS was refined to be more robot-tailored [8, 21].

Another study [16] claims that CVSS's range for vulnerability scoring is too short. CVSS focuses on vulnerabilities and is not able to evaluate attacks requiring the exploitation of multiple vulnerabilities (referred to as kill chain [10]). Therefore, the criticality of identified vulnerabilities is measured using a security metric based on an attack graph instead of including the dimension of

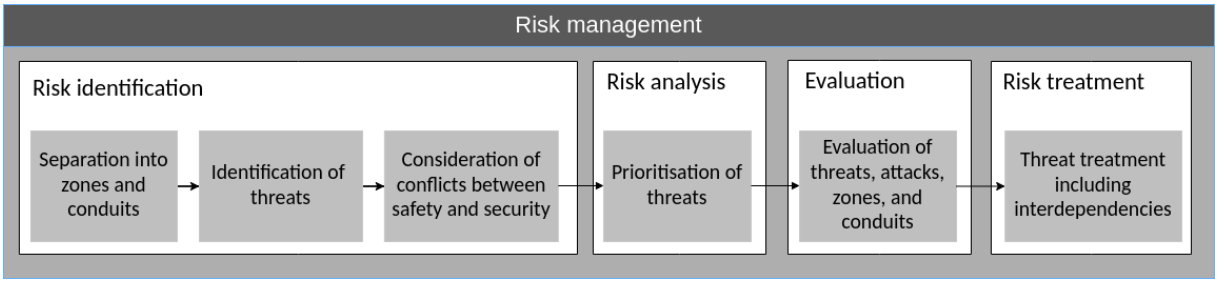

Figure 1: Threat model approach based on [9]

prior exploitation of other vulnerabilities needed. On one hand, this approach indeed increases the accuracy of evaluating the identified security risks and their prioritization to be addressed by the asset owner. However, the suggested evaluation does not provide qualitative information about the vulnerability's criticality. Therefore, an asset owner has no methodological suggestion on how many prioritized risks should be mitigated.

An evaluation of CVSS's base score was performed using Vulnerability Reward Programs (VRPs) [24]. It shows that there is a significant correlation between the CVSS base score and the VRP's severity ratings. This evaluation concluded that CVSS may still help to prioritize vulnerabilities.

2.2.2 Security Level.

The IEC 62443 standard series [11] defines four SLs levels that describe the skill level and resources needed by a threat actor to launch specific attacks. The higher the SL the more sophisticated the attack gets resulting in a higher skill-set and resources needed to perform the attack. Figure 2 illustrates this definition.

4 Security Level (SL)	
SL 1	Protection against **casual or coincidental** violation
SL 2	Protection against **intentional violation** using **simple means** with low resources, generic skills and low motivation
SL 3	Protection against intentional violation using **sophisticated means** with **moderate resources**, IACS specific skills and moderate motivation
SL 4	Protection against intentional violation using sophisticated means with **extended resources**, IACS specific skills and high motivation

Figure 2: SLs based on [11]

2.2.3 Relation between CVSS and SL.

The metric system CVSS consists of several metric groups and metrics, as Figure 3 illustrates. The *Base metric group* evaluates constant characteristics of a vulnerability over time that are independent of the user environment. This metric group may be divided into the Exploitability and Impact metrics. The *Exploitability metrics* weight the difficulty and technical means (e.g., computing resources) needed to exploit the vulnerability. The *Impact* sub-group measures the consequences of successful exploitation of the vulnerability. The *Temporal metric group* addresses changing characteristics of a vulnerability over time and is independent of the user environment.

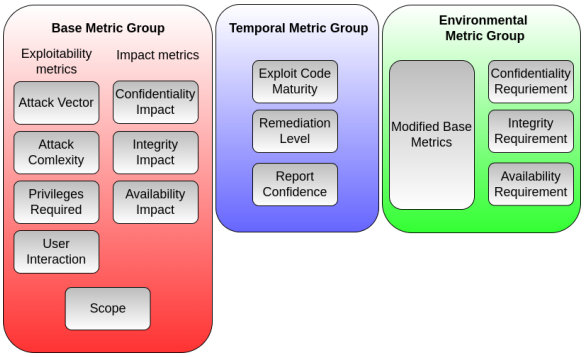

Figure 3: CVSS metric groups based on [5]

For instance, a ready-to-use exploit tool kit leads to a higher CVSS score. In contrast, the existence of an official path to mitigate the corresponding vulnerability leads to a lower CVSS score. The *Environmental metric group* evaluates a vulnerability's characteristics that depend on the user environment. For instance, installed security controls mitigate the consequences of a successful attack. The values provided in the Environmental metric group override their corresponding values of the Basic metric group (e.g., a defined MAV value overrides the prior defined AV value). The base metric group is mandatory, whereas the temporal and the environmental metric group are optional for calculation of the CVSS score. Based on the metrics CVSS and SL, the difficulty of the attack, the threat actor type, and the identified attacks could be mapped to an SL. CVSS and SL have an inverted relationship. CVSS rates critical attacks, attacks with a high probability and severe impact, with a high CVSS score. SL defines the needed security controls countering those attacks. The higher the SL, the more security controls are required to protect against a wider variety of attacks. This also means that a complex attack is rated with a low CVSS score while the very same attack is addressed with a high SL [9].

2.2.4 Safety Integrity Level.

The standard EN ISO 13849 [13] provides safety requirements and guidance on the principles for the design and integration of Safety-Related Parts of Control Systems (SRP/CS), including the design of software. Furthermore, it specifies characteristics that include the Performance Level (PL) required for carrying out safety functions. It applies to SRP/CS for high demand and continuous mode, regardless of the type of technology and energy used (e.g., electrical, hydraulic,

pneumatic, and mechanical) for all kinds of machinery. Examples of products that are parts of SRP/CS are relays, solenoid valves, position switches, Programmable Logic Controllers (PLCs), motor control units, two-hand control devices, and pressure-sensitive equipment.

Before applying this standard to a specific OT component, a safety risk assessment has to be done first, which may be achieved according to ISO 12100 [12]. This standard describes procedures for identifying hazards, estimating and evaluating risks during relevant phases of the machine life cycle, and eliminating hazards or sufficient risk reduction.

After performing a risk assessment according to ISO 12100 [12], the identified safety risks are evaluated using EN ISO 13849 [13]. Each identified risk is assigned with a PL that may be mapped to a SIL as Figure 4 demonstrates.

PL (Performance Level)	PFH_D (Probability of Dangerous Failure per Hour)	SIL
a	$\geq 10^{-5}$ to $<10^{-4}$	None
b	$\geq 3 \times 10^{-6}$ to $<10^{-5}$	1
c	$\geq 10^{-6}$ to $<3 \times 10^{-6}$	1
d	$\geq 10^{-7}$ to $<10^{-6}$	2
e	$\geq 10^{-8}$ to $<10^{-7}$	3

Figure 4: Relation between PL and SIL based on [13]

3 PROPOSED RISK EVALUATION

The proposed risk evaluation is a key contribution of this work. We propose to evaluate the identified attacks using the metric CVSS.

3.1 Mapping CVSS to SL

The basic idea of the CVSS mapping is the following: A low CVSS score leads to an high SL and the other way around. SL considers the attacker's intention, motivation, skills, and resources but not the potential impact caused by the attack. Therefore, the proposed mapping does not consider the impact metrics of the CVSS base and environmental score (cf. Figure 3). Each metric value of the CVSS score has a dedicated numerical value assigned. Since CVSS needs to have at least one impact metric value rated higher than "None" to generate a CVSS score other than "0.0", the impact metric availability (A) got the value "Low" assigned to perform the mappings. A numerical value of 1.6 results when the least possible values for the Exploitability metrics of the base score, the temporal score, and the Exploitability metrics of the environmental score are added. Adding the highest possible values for the same metrics of the CVSS score leads to a numerical value of 5.8. Thus the score used for the mapping ranges between 1.6 and 5.8. This range is divided into four sub-ranges that define the mapping, as Table 1 lists. A stakeholder analysis [8] showed that SL-T 4 is very unlikely to be used in the industry. Therefore, a corresponding short range of CVSS mapping score is used that results in SL 4. The suggested ranges for mapping CVSS to SL were validated by the application to identified attacks, where Table 3 shows a subset these attacks.

Table 1: Ranges of CVSS to SL mapping

CVSS Mapping score range	SLs
4.5 - 5.8	SLs 1
3.1 - 4.4	SLs 2
1.7 - 3.0	SLs 3
1.6 - 1.7	SLs 4

3.2 Prioritisation of measures for risk treatment

The suggested prioritisation of the countermeasures for risk treatment is based on the following indicators, where the first three represent probability and the remaining three impact of the corresponding risk:

- Probability
  - Threat actor
  - Calculated SL of attack
  - Threats enabling multiple identified attacks
- Impact
  - SIL of OT component
  - Business-criticality of the asset
  - SL-T of zone targeted by an attack

The indicator "Threat actor" defines the skill level and resources needed to launch successful attacks based on the definition in Section 2.1. The threat actors are mapped to an SL by linking the definition of threat actors [3, 18] to a dedicated SL according to the definition of the IEC 62443 standard series [11], as Table 2 lists.

Table 2: Threat actors assigned to SL

Threat actor	SLs
Basic user	SLs 1
Insider	SLs 2
Hacktivist	SLs 2
Cyber criminal	SLs 3
Nation-state attacker	SLs 4
Terrorist	SLs 4

The indicator "Calculated SL of attack" is described in Section 3.1. "Threats enabling multiple identified attacks" are identified during the *risk analysis* phase of the used threat model approach [9], where attacks are modeled based on identified threats in the system architecture (cf. Section 5.1). Therefore, two different SL values are used for this prioritization: One resulting from the difficulty for the attack and another one defining the skill level of the attacker. Using two SLs from an attack's and a threat actor's viewpoint respectively considers sophisticated attacks that may not be able to be launched when the required skill level and resources are not reached. "SIL of OT component" considers the SIL of an effected safety-relevant OT component as an impact-indicator of the risk. The "Business-criticality of the asset" rates the potential impact of an asset when compromised. For instance, a successful attack against a domain controller results in more serious consequences than an attack against an HMI. "SL-T of zone targeted by attack"

defines the desired level of security requirements of a zone. A countermeasure is recommended when the attack is more advanced than the requirements of the corresponding zone's desired SL-T addresses. The prioritization of the attacks is evaluated according to Equation (1) using the following definition of variables:

- z is the effected zone or conduit of the attack
- a is the SL of the attack
- t is the SL of the threat actor of the attack
- s is the SIL of the attacked OT component, if applicable
- c is the business-criticality of the attacked OT component
- m is the number of attacks caused by a underlying threat
- n is the number of attacks identified

$$p(x) = \begin{cases} \min(a_0 \dots a_n) + \min(t_0 \dots t_n) + \max(s_0 \dots s_n) \\ \quad + \max(c_0 \dots c_n) + \max(m_0 \dots m_n) & \text{if } x_z \geq x_a \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

4 USE CASE

A stakeholder analysis was performed and published in a previous work [8]. Vendors of OT components, integrators, and asset owners of industrial architectures contributed to this work. This stakeholder analysis provides insights about common characteristics and application practices in OT environments from the perspective of each party of the supply chain. The following topics were considered in this study:

- OT components
- Communication and integration
- Operation and change management
- Security countermeasures in industrial architectures
- Safety/security conflicts and interdependencies
- Risk management

The results of the stakeholder analysis were used to derive a use case as Figure 5 illustrates. This modeled use case serves as a SUC for applying the threat modeling approach and conducting the proposed risk assessment methodology. This use case was created according to PERA [22], showing each level on the left side of the figure. Furthermore, the modeled industrial architecture was separated into zones and conduits, including the assignment of an specific SL-T for each zone and conduit, as the standard IEC 62443 [11] suggests. Since the traditional office IT is out of scope of IEC 62443, the management level was not considered during the risk assessment of this use case.

5 APPLICATION AND RESULTS

This section shows the results of applying the threat modeling approach [9] to the use case illustrated in Figure 5.

5.1 Risk analysis

Applying the threat modeling approach [9] to the use case described in Section 4 leads to the results listed in Table 3. The presented results are an excerpt of all the identified attacks. Since the *risk identification* was discussed in detail in [4] this paper focuses on *risk analysis* and *risk evaluation* instead. Instead, Table 3 lists the results of the *risk analysis* and is defined as follows:

- A-ID: ID of the identified attack
- Threat Event: Description of the identified attack
- Threat Actor: Threat actor with the least skill level and resources applicable
- Attack Sequence: Vulnerabilities exploited to launch the attack, based on the *risk identification* phase
- Interdependencies addressed: Effects to the relation between security and safety functions

Script kiddies may perform Denial of Service (DoS) attacks against the Internet-facing web server, leading to no safety-relevant consequences.

Insiders may exploit a buffer overflow vulnerability to modify safety parameters. A threat in the attack sequence modifies safety parameters and interferes with a safety function, thus addressing the dependency of safety on security. These threat actors may also launch DoS attacks on a robot to sabotage the production without needing to circumvent physical security measures.

A hacktivist may extract or destroy critical operational data via exploiting an SQL injection vulnerability at the data historian after entering the internal network by a successful phishing attempt leading to financial but no safety-relevant consequences. They may jam or intercept radio-based signals (e.g., wirelessHART) used in the field for wireless transmission of sensor data.

Cyber criminals may enter the internal network via a successful phishing attempt to introduce a ransomware infection that could lead to safety-relevant consequences. These threat actors may also establish a command & control (C2) server to disrupt physical operations remotely using a variant of social engineering by placing malicious USB sticks at the target's site. In this case, the threats enabling to modify and update the application logic of the targeted PLC show another dependency on safety from security. They may compromise the domain controller at the management level, which gets replicated over time automatically to the domain controller placed at the supervisory level to gain access to the targeted SUC. This enables access to all OT components using the domain controller for centralized user management. Afterwards, the attacker can access and modify the safety parameters of the cobot via the web interface [7]. A cyber criminal may exploit the remote maintenance access via session hijacking to directly access the robot to be maintained and modify its safety parameters.

Nation-state actors or terrorists may manage to access restricted physical areas via the emergency procedure (e.g., triggered by pushing an emergency stop button) and gain access to production lines. This attack is an example of the dependency of security on a safety function since triggering a safety function enabled unauthorized physical access. These threat actors may continue this attack with the sabotage of emergency buttons connected to safety-relevant OT components, such as PLCs, robots, or the cobot. As the malware TRITON [15] demonstrated, attackers may corrupt an SIS remotely. This attack impacts the safety function that may impact security again when the safety function of the SIS is executed delayed leading to availability loss due to the execution of the safety function.

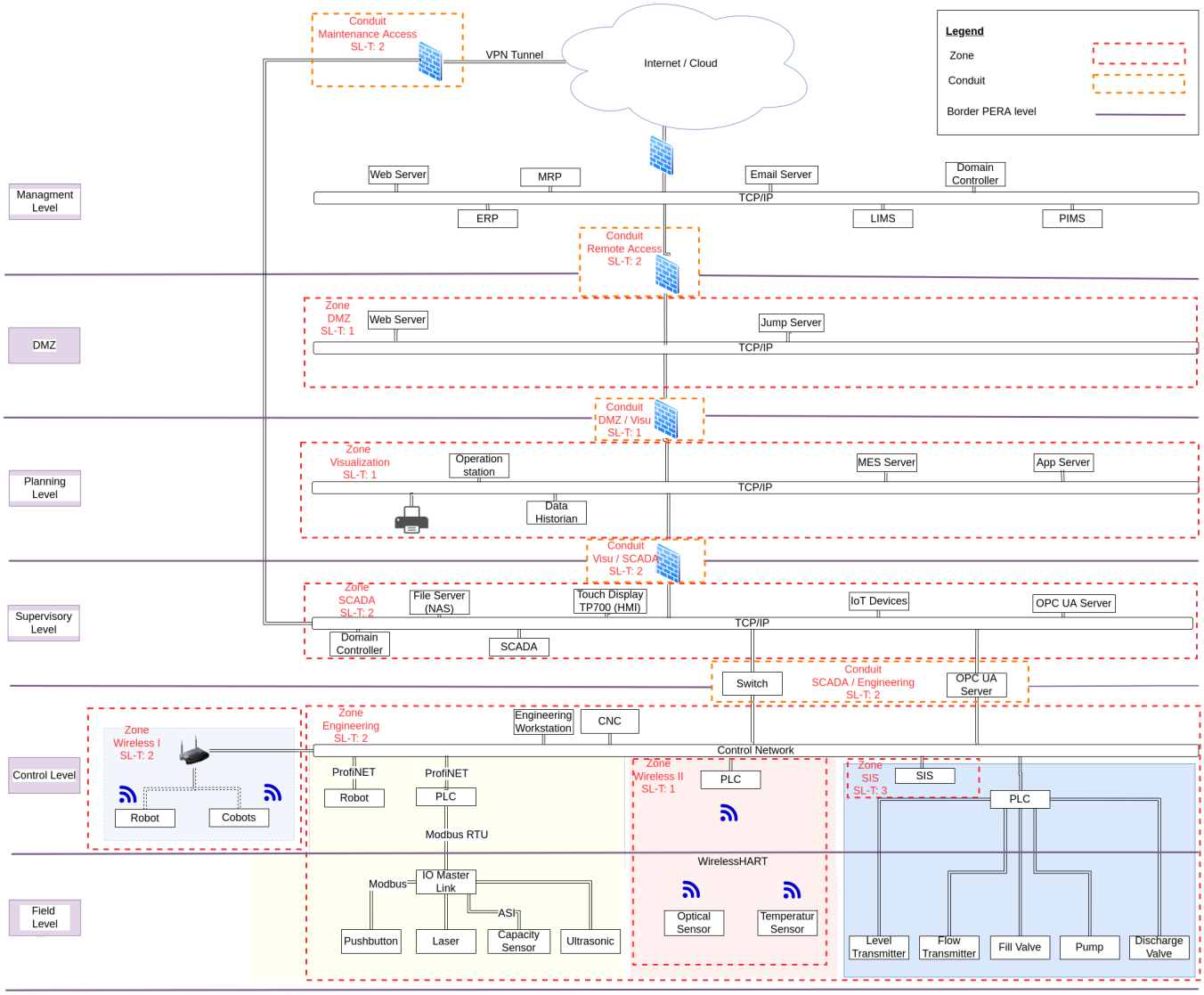

Figure 5: Use case based on a stakeholder analysis [8]

5.2 Evaluation

Table 4 shows the results when applying the proposed evaluation scheme to the attack listed in Table 3. The columns of this table are defined as follows:

- A-ID: ID of the identified attack
- CVSS vector: The overall CVSS vector of the identified attack
- CVSS base score: The overall CVSS base score of the identified attack
- Mapped score: Resulting score after performing the mapping according to Table 1.
- Attack SL: Classification of the skill level and resources needed to successfully launch this attack. This value is a result of the mapping introduces in Table 1.
- Threat actor SL: Minimum SL of a threat actor who is able to launch this attack successfully according to Table 2.
- OT component SIL: If a safety-relevant OT component is affected during the attack, its SIL increases the potential impact of the resulting risk of the attack.
- Priority neglecting defined SL-T: Resulting priority considering the condition $x_z < x_a$ (cf. Equation (1)).
- Priority considering defined SL-T: Resulting priority including considering the condition $x_z \geq x_a$ (cf. Equation (1)).

From the attacks listed in Table 3, the highest priority is on the protection against the corruption of safety parameters from a compromised user of the domain controller located in the management level that was synced to the domain controller located in the supervisory level. Ransomware infection introduced through a phishing

Table 3: Excerpt of identified attacks using [9]

A-ID	Threat Event	Threat Actor	Attack Sequence	Interdepdencies addressed
A1	Buffer overflow vulnerability on cobot's firmware gets exploited	Insider	1) exploitation firmware vulnerability 2) modify safety parameters	Threat impacts safety function
A2	Ransomware infection	Cyber criminal	1) phishing e-mail 2) accessible file server gets infected 3) infection spreads to users accessing the file server	-
A3	External party established C2 server and disrupt physical operations remote	Cyber criminal	1) social engineering (placing USB sticks) 2) application logic of PLC gets modified 3) update modified logic	Threat impacts safety function
A4	DoS attack against robot to sabotage production	Insider	1) attacker connects to control network 2) exploit DoS vulnerability	Threat impacts safety function
A5	Compromised Active Directory (AD) impacts engineering station that influences physical operations	Cyber criminal	1) compromized AD from management level gets replicated to supervisory level 2) privilege escalation on engineering workstation 3) corrupt user compromized safety parameters via GUI of cobot (cf. [7])	Threat impacts safety function
A6	Extraction and/or destruction of critical operational data	Hacktivist	1) phishing 2) privilege escalation of operator station 3) exploitation of SQL injection vulnerability at data historian	-
A7	Direct access to robot via remote maintenance access	Cyber criminal	1) remote session hijacking 2) access robot via default credentials 3) modify safety parameters	Threat impacts safety function
A8	Jamming/intercepting of wireless signals	Hacktivist	1) social engineering (e.g., disguise as interal or 3rd party employee) 2) jamming wirelessHART signals	-
A9	Access restricted physical areas via emergency procedure	Nation-state actor/terrorist	1) social engineering (e.g., disguise as interal or 3rd party employee) 2) gain physical access to production lines	Safety function introduces threat
A10	Sabotage of emergency button connected to PLCs/cobot/robot	Nation-state actor/terrorist	1) social engineering (e.g., disguise as interal or 3rd party employee) 2) gain physical access to production lines 3) destroy safety button / exchange with modified safety button	Safety function introduces threat
A11	Corruption of SIS (e.g., TRITON [15])	Nation-state actor/terrorist	1) compromized AD from management level gets replicated to supervisory level 2) privilege escalation on engineering workstation 3) install SIS-malware	Threat impacts safety function Safety function introduces threat
A12	DoS attack against web server	Basic user/script kiddie	1) DoS attack from the Internet	-

Table 4: Evaluation of identified attacks

A-ID	CVSS vector	CVSS base score	Mapped score	Attack SL	Threat actor SL	OT component SIL	Priority neglecting SL-T	Priority considering SL-T
A1	CVSS:3.1/AV:L/AC:H/PR:N/UI:N/S:C/C:H/I:H/A:H	8.1	3.2	2	2	2	7	4
A2	CVSS:3.1/AV:N/AC:L/PR:N/UI:R/S:C/C:L/I:N/A:N	9.3	4.7	1	3	-	2	2
A3	CVSS:3.1/AV:N/AC:H/PR:H/UI:R/S:C/C:L/I:N/A:N	7.6	2.6	3	3	2	3	0 (recommendation)
A4	CVSS:3.1/AV:L/AC:L/PR:H/UI:N/S:C/C:N/I:N/A:H	6.0	3.2	2	2	2	7	4
A5	CVSS:3.1/AV:A/AC:L/PR:L/UI:N/S:C/C:L/I:H/A:N	7.6	4.1	2	3	2	1	1
A6	CVSS:3.1/AV:A/AC:L/PR:L/UI:N/S:U/C:H/I:H/A:N	7.3	3.5	2	2	-	10	0 (recommendation)
A7	CVSS:3.1/AV:N/AC:H/PR:H/UI:R/S:C/C:H/I:H/A:H	7.6	2.6	3	3	2	6	0 (recommendation)
A8	CVSS:3.1/AV:P/AC:L/PR:N/UI:N/S:U/C:N/I:N/A:L	2.4	2.4	2	2	-	7	0 (recommendation)
A9	CVSS:3.1/AV:P/AC:H/PR:L/UI:R/S:C/C:L/I:N/A:N	1.9	1.9	3	4	1	10	0 (recommendation)
A10	CVSS:3.1/AV:P/AC:H/PR:H/UI:R/S:C/C:H/I:H/A:H	6.8	1.8	3	4	2	3	0 (recommendation)
A11	CVSS:3.1/AV:A/AC:H/PR:H/UI:N/S:C/C:H/I:H/A:H	7.6	2.6	3	3	3	3	3
A12	CVSS:3.1/AV:N/AC:H/PR:N/UI:N/S:U/C:N/I:N/A:L	3.7	5.8	1	1	-	12	6

e-mail targeting the file server in the SCADA zone (cf. Figure 5) to spread to users connected to this server has the second-highest priority assigned. The remote corruption of an SIS using a similar entry point to the SUC, namely corrupting a user managed by the domain controller in the management level that gets replicated to the domain controller in the supervisory level, has the third priority level considering the countermeasures. Insiders exploiting an easy-to-exploit legacy firmware vulnerability due to patch management constraints in OT environments [8] and performing DoS attacks against the robot located in the engineering zone are risks with the fourth priority assigned. A remote DoS attack against the web server located in the DMZ zone has the most minor priority level due to its low impact but high probability. The protection against the remaining attacks identified is considered as a recommendation without priority since the defined SL-T of the zone the attack is targeted lower than the SL of the attack, reflecting the identified typical risk appetite of an asset owner for similar zones [8].

6 DISCUSSION

The proposed risk evaluation methodology provides multiple benefits. The prioritization of the analyzed risks considers various security- and safety-relevant system-specific factors to suggest risk treatment. This replaces vague indicators of industrial security like potential financial, reputation, or compliance loss. The introduced evaluation methodology is compatible with several existing risk assessment approaches due to the usage of commonly known metrics. Additionally, the results of other risk identifications or assessments may be included in the proposed approach. For instance, the results of a penetration test may be included since the identified vulnerabilities and attacks are typically rated with CVSS [8]. Another example is that the verified SIL of an OT component may be used as an impact indicator during the risk evaluation. Instead of viewing security and safety isolated, the introduced risk evaluation scheme provides a combined view of both protection goals, including their interdependencies. The presented mapping of CVSS to SL does intentionally only consider the probability sub-metrics of CVSS since the skill level and resources needed to launch an attack

(which define an SL) are not necessarily linked to the impact caused. Otherwise, the calculation would lead to higher SLs of the attacks, which are less likely to meet the defined SL-T of a zone or conduit.

It was discovered during the application that the used threat modeling approach [9] could be optimized by merging the definition of zones and conduits with the definition of trust boundaries during the *risk identification* phase instead of doing this task separately in the *risk treatment*, as Figure 1 shows.

The results of this work demonstrate meeting the requirements *attack modeling*, *risk analysis*, *common metrics*, and *efficiency* on threat model approaches [9].

7 CONCLUSION AND FUTURE WORK

This work presents a risk assessment methodology for OT environments and focuses on risk evaluation. Metrics commonly used in the industry were considered in the evaluation, which increases the compatibility with other existing risk assessment approaches and the acceptance in the industry since the resources needed to implement this schema are low as no mapping or parsing to self-defined metrics is required. The introduced risk assessment methodology was applied to a sample industrial architecture derived from a stakeholder analysis to further address the industry's acceptance.

Future work will formalize the identified characteristics and requirements of industrial architectures in a system model of OT environments. This system model will include safety and security requirements while considering interdependencies and potential conflicts between safety and security. Therefore, combining the upcoming system model with the presented threat modeling approach enables the application of the approach to arbitrary industrial architectures. Developing the system model as an ontology (e.g., using Web Ontology Language (OWL)) introduces a semi-automatized possibility to describe OT environments and perform system-tailored risk analysis based on identified system-specific threats and attacks. Furthermore, an OT protection catalog will be created considering safety, security, and their interdependencies. If a conflict between safety and security requirements arises in the SUC, the OT protection catalog will suggest a system-tailored set of technical and organizational countermeasures addressing this conflict to fulfill both security and safety requirements.

ACKNOWLEDGMENTS

This work was enabled by TÜV AUSTRIA #SafeSecLab Research Lab for Safety and Security in Industry, a research cooperation between TU Wien and TÜV AUSTRIA.

REFERENCES

[1] Sadek Rayan Aktouche, Mohamed Sallak, Abdelmadjid Bouabdallah, and Walter Schön. 2021. Towards Reconciling Safety and Security Risk Analysis Processes in Railway Remote Driving. In *2021 5th International Conference on System Reliability and Safety (ICSRS)*. 148–154. https://doi.org/10.1109/ICSRS53853.2021.9660764

[2] Blake E. Strom and Joseph A. Battaglia and Michael S. Kemmerer and William Kupersanin and Douglas P. Miller and Craig Wampler and Sean M. Whitley and Ross D. Wolf. 2017. *Finding Cyber Threats with ATT&CK-Based Analytics*. Technical Report. The MITRE Corporation.

[3] Clint Bodungen, Bryan Singer, Aaron Shbeeb, Kyle Wilhoit, and Stephen Hilt. 2016. *Hacking Exposed Industrial Control Systems: ICS and SCADA Security Secrets & Solutions* (1 ed.). McGraw-Hill Education, New York. 544 pages. https://doi.org/10.1036/9781259589720

[4] Patrick Denzler, Siegfried Hollerer, Thomas Frühwirth, and Wolfgang Kastner. 2021. Identification of security threats, safety hazards, and interdependencies in industrial edge computing. In *2021 IEEE/ACM Symposium on Edge Computing (SEC)*. 397–402. https://doi.org/10.1145/3453142.3493508

[5] FIRST (FIRST.Org, Inc.). 2019. Common Vulnerability Scoring System version 3.1, Specification Document, Revision 1.

[6] Marcus Geiger, Jochen Bauer, Michael Masuch, and Jörg Franke. 2020. An Analysis of Black Energy 3, Crashoverride, and Trisis, Three Malware Approaches Targeting Operational Technology Systems. In *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Vol. 1. 1537–1543. https://doi.org/10.1109/ETFA46521.2020.9212128

[7] Siegfried Hollerer, Clara Fischer, Bernhard Brenner, Maximilian Papa, Sebastian Schlund, Wolfgang Kastner, Joachim Fabini, and Tanja Zseby. 2021. Cobot attack: a security assessment exemplified by a pecific collaborative robot. *Procedia Manufacturing* 54 (2021), 191–196. https://doi.org/10.1016/j.promfg.2021.07.029 10th CIRP Sponsored Conference on Digital Enterprise Technologies (DET 2020) – Digital Technologies as Enablers of Industrial Competitiveness and Sustainability.

[8] Siegfried Hollerer, Wolfgang Kastner, and Thilo Sauter. 2021. Safety und Security - ein Spannungsfeld in der industriellen Praxis. *e & i Elektrotechnik und Informationstechnik* 138, 449–453. https://doi.org/10.1007/s00502-021-00930-0

[9] Siegfried Hollerer, Wolfgang Kastner, and Thilo Sauter. 2021. Towards a Threat Modeling Approach Addressing Security and Safety in OT Environments. In *2021 17th IEEE International Conference on Factory Communication Systems (WFCS)*. 37–40. https://doi.org/10.1109/WFCS46889.2021.9483591

[10] Eric Hutchins, Michael Cloppert, and Rohan Amin. 2011. Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains. *Leading Issues in Information Warfare & Security Research* 1 (2011).

[11] International Electrotechnical Commission (IEC). 2020. IEC 62443, Security for industrial automation and control systems.

[12] International Organization for Standardization (ISO). 2010. ISO 12100 Safety of machinery — General principles for design — Risk assessment and risk reduction.

[13] International Organization for Standardization (ISO). 2015. ISO 13849 Safety of machinery — Safety-related parts of control systems.

[14] Juergen Jasperneite, Thilo Sauter, and Martin Wollschlaeger. 2020. Why We Need Automation Models: Handling Complexity in Industry 4.0 and the Internet of Things. *IEEE Industrial Electronics Magazine* 14, 1 (2020), 29–40. https://doi.org/10.1109/MIE.2019.2947119

[15] Jin-woo Myung ; Sunghyuck Hong. 2019. ICS malware Triton attack and countermeasures.. In *International Journal of Emerging Multidisciplinary Research*.

[16] Marjan Keramati. 2017. A novel system for quantifying the danger degree of computer network attacks. In *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*. https://doi.org/10.1109/KBEI.2017.8324906

[17] Yazid Merah and Tayeb Kenaza. 2021. Ontology-Based Cyber Risk Monitoring Using Cyber Threat Intelligence. In *The 16th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES 2021)*. Association for Computing Machinery, New York, NY, USA, Article 88, 8 pages. https://doi.org/10.1145/3465481.3470024

[18] Marco Rocchetto and Nils Ole Tippenhauer. 2016. On Attacker Models and Profiles for Cyber-Physical Systems. In *Computer Security – ESORICS 2016*, Ioannis Askoxylakis, Sotiris Ioannidis, Sokratis Katsikas, and Catherine Meadows (Eds.). Springer International Publishing, Cham, 427–449.

[19] Jonathan Spring, Eric Hatleback, Allen Householder, Art Manion, and Deana Shick. 2021. Time to Change the CVSS? *IEEE Security Privacy* 19, 2 (2021), 74–78. https://doi.org/10.1109/MSEC.2020.3044475

[20] Max van Haastrecht, Injy Sarhan, Alireza Shojaifar, Louis Baumgartner, Wissam Mallouli, and Marco Spruit. 2021. A Threat-Based Cybersecurity Risk Assessment Approach Addressing SME Needs. In *The 16th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES 2021)*. Association for Computing Machinery, New York, NY, USA, Article 158, 12 pages. https://doi.org/10.1145/3465481.3469199

[21] Víctor Mayoral Vilches, Endika Gil-Uriarte, Irati Zamalloa Ugarte, Gorka Olalde Mendia, Rodrigo Izquierdo Pisón, Laura Alzola Kirschgens, Asier Bilbao Calvo, Alejandro Hernández Cordero, Lucas Apa, and César Cerrudo. 2021. Towards an open standard for assessing the severity of robot security vulnerabilities, the Robot Vulnerability Scoring System (RVSS). arXiv:1807.10357 [cs.RO]

[22] Martin Wollschlaeger, Thilo Sauter, and Juergen Jasperneite. 2017. The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0. *IEEE Industrial Electronics Magazine* 11, 1 (2017), 17–27. https://doi.org/10.1109/MIE.2017.2649104

[23] Shengwei Yi, Hongwei Wang, Yangyang Ma, Feng Xie, Puhan Zhang, and Liqing Di. 2018. A Safety-Security Assessment Approach for Communication-Based Train Control (CBTC) Systems Based on the Extended Fault Tree. In *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. 1–5. https://doi.org/10.1109/ICCCN.2018.8487464

[24] Awad Younis, Yashwant K. Malaiya, and Indrajit Ray. 2016. Evaluating CVSS Base Score Using Vulnerability Rewards Programs. In *ICT Systems Security and Privacy Protection*, Jaap-Henk Hoepman and Stefan Katzenbeisser (Eds.). Springer International Publishing, Cham, 62–75.