Enhancing the Resilience of Cyber-Physical Systems by Protecting the Physical-World Interface

Rainer Falk, Steffen Fries
Corporate Technology
Siemens AG
Munich, Germany
e-mail: {rainer.falk|steffen.fries}@siemens.com

***Abstract*—Cyber physical systems operate and supervise physical, technical systems using information and communication technology, also called Operation Technology (OT). Cyber security solutions focus on the OT part, i.e., on the information and communication technology. The focus of cyber security is protection, detection, and respondence to cyber attacks. Cyber resilience aims at delivering an intended outcome despite attacks and adverse cyber events and even failures not directly caused by attacks. Protecting the link between the control systems and the physical world has been addressed only in some very specific cases, e.g., charging of electric vehicles. We propose a physical-world firewall that limits the impact on the physical world of a successful attack of automation systems, thereby enhancing the resilience of cyber-physical system against successful attacks against software-based functionality of its OT systems.**

Keywords*—*cyber security; cyber resilience; system integrity; cyber physical systems; industrial automation and control system; Internet of Things.

I. INTRODUCTION

The common focus of IT security relates to IT-based control equipment and data communication, using e.g., Ethernet, wireless LAN (WLAN), and Internet protocol (IP) communication. In addition to this, in OT systems, also the field level comprising sensors and actuators connected to the Operation Technology (OT) automation and control system has to be considered down to the interface between the control system and the physical world via sensors and actuators.

Traditionally, IT security has been focusing on information security, protecting confidentiality, integrity, and availability of data at rest and data in transit, and sometimes also protecting data in use by confidential computation. In Cyber-Physical Systems (CPS), major protection goals are availability, meaning that automation systems stay productive, and system integrity, ensuring that it is operating as intended. Typical application domains are factory automation, process automation, building automation, railway signaling systems, and power system management. Cyber security is covering different phases during operation as there are protect, detect, and react: Protecting against threats, detecting when an attack has occurred, and recovering from attacks.

We see resilience of cyber-physical systems as an important further protection goal, to limit the effect of potential successful attacks on a cyber-physical system in the physical world. In addition, resilience also addresses system stability to cope with failure scenarios not caused by a successful attack. It can be rather seen as a strategy than a specific technology. Our objective is to increase the robustness with respect to intentional attacks, although resilience in general would consider also accidental failures. This paper, being an extended version of [1], puts the focus on the interface between the OT system, i.e., the automation and control system, and the physical world, proposing an additional layer of defense for cyber physical systems. It can be considered as "physical world firewall", limiting the access to the physical world by the OT system.

After giving an overview on cyber physical systems and on industrial cyber security in Sections II and III, a new approach on protecting the interface of a CPS between the cyber-world and the physical world is described in Section IV. It is a concept to increase the resilience of a CPS when being under attack. Aspects to evaluate the new approach are discussed in Section V. Section VI concludes the paper.

II. CYBER PHYSICAL SYSTEMS

A cyber-physical system, e.g., an industrial automation and control system, monitors and controls a technical system. Examples are process automation, machine control, energy automation, and cloud robotics. Automation control equipment with sensors (S) and actuators (A) is connected directly with automation components, or via remote input/output modules. The technical process is controlled by measuring its current state using the sensors, and by determining the corresponding actuator signals.

Figure 1 shows an example of an industrial automation and control system, comprising different control networks connected to a plant network and a cloud backend system. Separation of the network is typically used to realize distinct control networks with strict real-time requirements for the interaction between sensors and actuators of a production cell, or to enforce a specific security policy within a production cell.

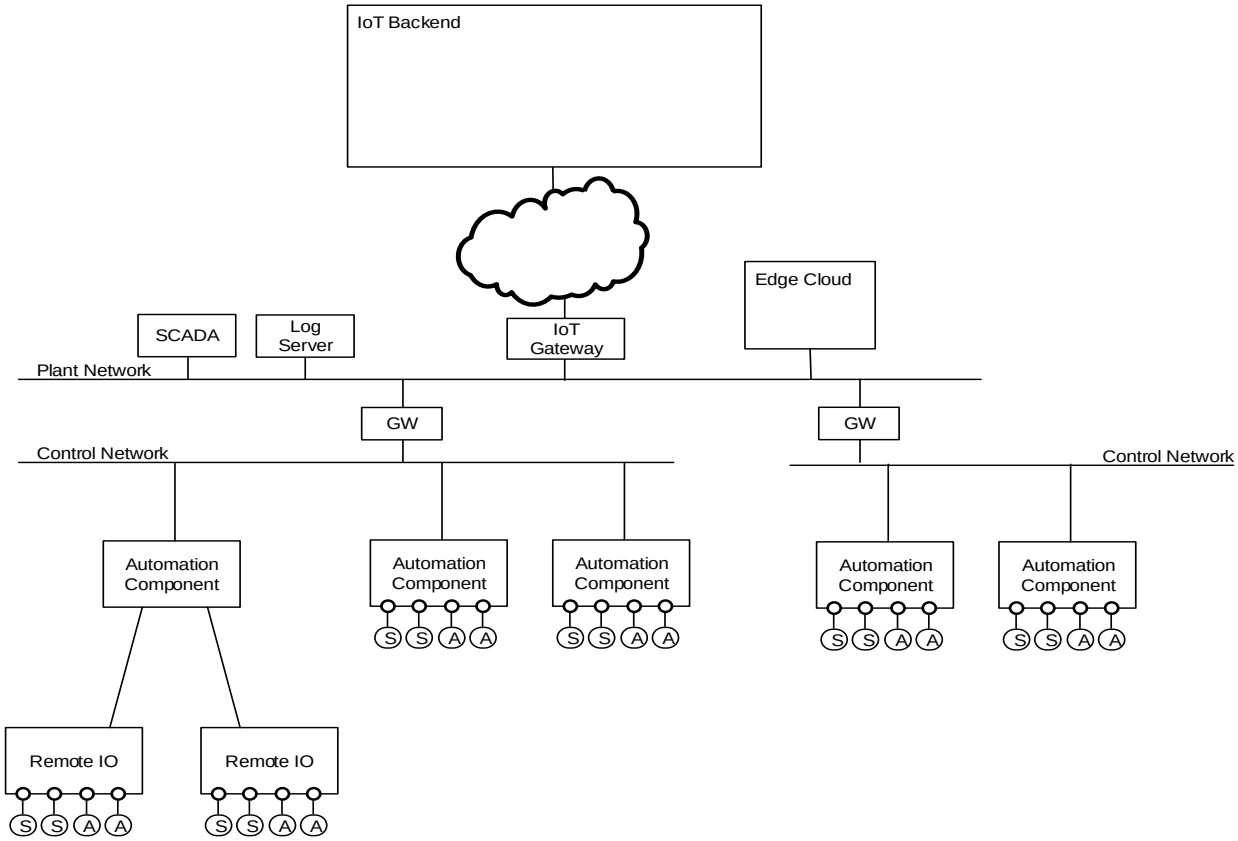

Figure 1. Example CPS System

Such an industrial automation and control system is an example of a cyber-physical system and are utilized in various automation domains, including discrete automation (factory automation), process automation, railway automation, energy automation, and building automation.

Figure 2 shows the typical structure of automation components. The functionality realized by an automation component is largely defined by the firmware/software and the configuration data stored in its flash memory.

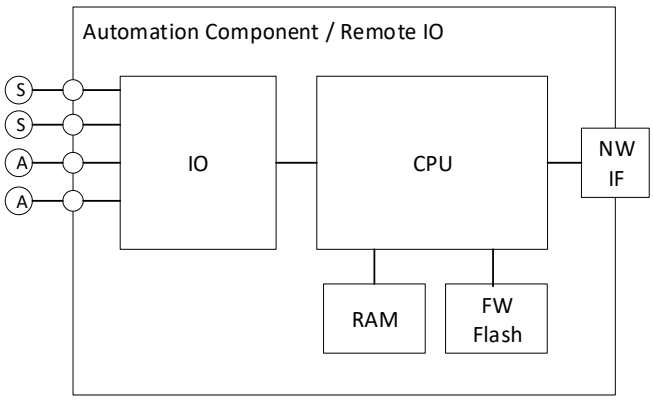

Figure 2. Automation Component

In practice, it has to be assumed that each software component may comprise vulnerabilities, independent of the effort spend to ensure high software quality. This is a reason why automation systems are usually organized in separate security zones. Network traffic can be filtered using network firewalls between different zones, limiting the impact of an impact in one security zone on other connected security zones. In addition, it is often not possible to fix known vulnerabilities immediately by installing a software update, as updates have to be tested thoroughly in a test system before being installed in an operational system, and as an installation is often possible only during a scheduled maintenance window. Also, the priorities of security objectives in different security zones are often different.

In cyber physical systems, the impact of a vulnerability in the OT system may not only affect data and data processing as in classical IT, but it may have an effect also on the physical world. For example, production equipment could be damaged, or the physical process may operate outside the designed physical boundaries, so that the produced goods may not have the expected quality.

III. INDUSTRIAL CYBER SECURITY

Protecting industrial automation control systems against intentional attacks is increasingly demanded by operators to ensure a reliable operation, and meanwhile also by regulation.

This section gives an overview on industrial security, and on the main relevant industrial security standard IEC 62443 [11] and integrity security requirements.

A. Industrial CPS Security Requirements

Industrial security is called also Operation Technology security (OT security), to distinguish it from general Information Technology (IT) security. Industrial systems have not only different security requirements compared to general IT systems, but come also with specific side conditions that prevent that security concepts established in the IT domain can be directly applied in an OT environment. For example, availability and integrity of an automation system often have a higher priority than confidentiality. As an example, high availability requirements, different organization processes (e.g., yearly maintenance windows), and required certifications may prevent the immediate installations of updates.

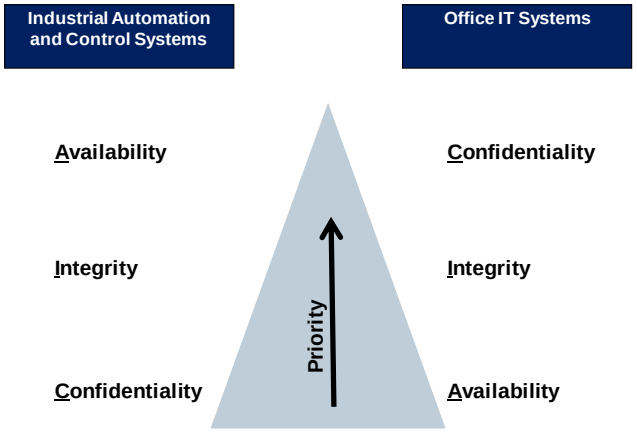

Figure 3. The CIA Pyramid [9]

The three basic security requirements are confidentiality, integrity, and availability. They are also named "CIA" requirements. Figure 3 shows that in common IT systems, the priority is "CIA". However, in automation systems or industrial IT, the priorities are commonly just the other way around: Availability has typically the highest priority, followed by integrity. Confidentiality is often no strong requirement for control communication, but may be needed to protect critical business know-how. As shown graphically, the CIA pyramid is inverted (turned upside down) in many automation systems.

Specific requirements and side conditions of industrial automation systems like high availability, planned configuration (engineering info), long life cycles, unattended operation, real-time operation, and communication, as well as safety requirements have to be considered when designing a security solution. Note that safety addresses undesired impacts originating from a technical system to the environment, e.g., in the case of a malfunction, while security addresses intentional attacks on the technical system. Often, an important aspect is that the applied security measures do not put availability and integrity of the automation system at risk. Depending on the considered industry (vertical), they may also be part of the critical infrastructure domain, for which security requirements are also imposed for instance by the European Network and Information Systems (NIS) directive [10] or country specific realizations of the directive. Further security requirements are provided by applying standards defining functional requirements, for instance defined in IEC 62443. The defined security requirements can be mapped to different automation domains, including energy automation, railway automation, building automation, process automation.

Security measures to address these requirements range from security processes, personal and physical security, device security, network security, and application security. No single security technology alone is adequate, but a combination of security measures addressing prevention, detection, and reaction to incidents is required ("defense in depth").

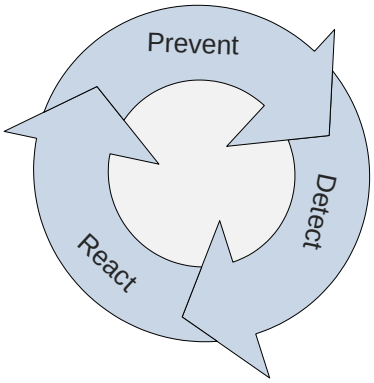

Figure 4. Prevent Detect React Cycle

Also, overall security has to address the areas prevent, detect, and react, see Figure 4. It is not sufficient to only define measures to protect against attacks. The capability has also foreseen to detect attacks, and to define measures to react adequately once an attack has been detected. The physical world firewall described in this paper is targeting the "react" phase, limiting the impact of a successful attack.

B. Overview IEC 62443 Industrial Security Standard

The international industrial security standard IEC 62443 [11] is a security requirements framework defined by the International Electrotechnical Commission (IEC). It addresses the need to design cybersecurity robustness and resilience into industrial automation and control systems, covering both organizational and technical aspects of security over the life cycle. Specific parts of this framework are applied successfully in different automation domains, including factory and process automation, railway automation, energy automation, and building automation. The standard specifies security for industrial automation and control systems (IACS) and covers both, organizational and technical aspects of security. Specifically addressed for the industrial domain is

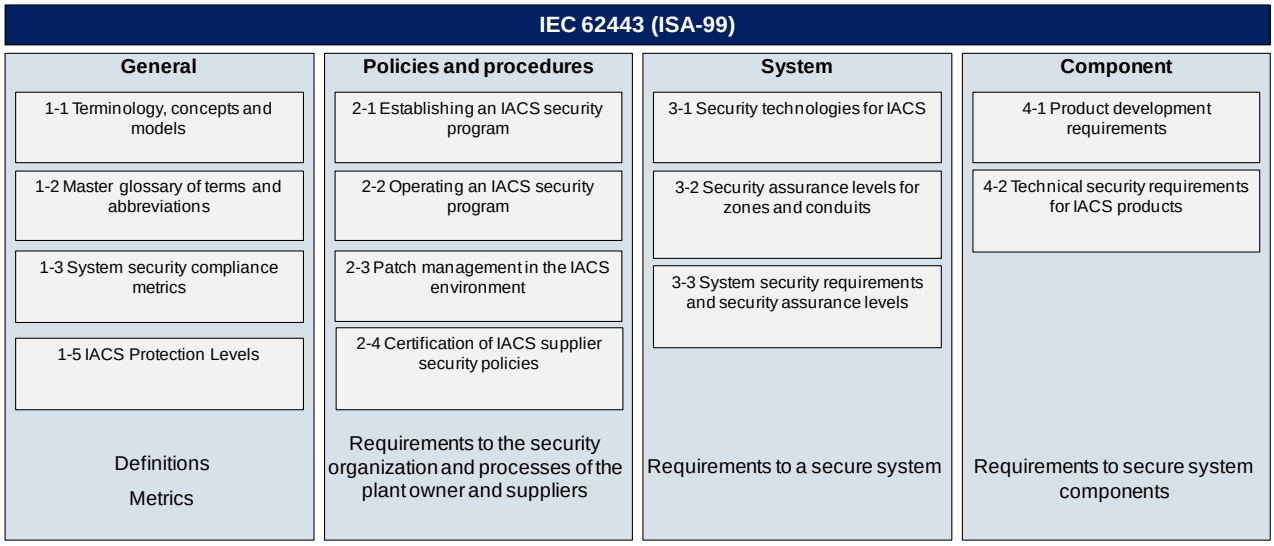

Figure 5. IEC 62443 Industrial Security Standard – Overview

the setup of a security organization and the definition of security processes as part of an information security management system (ISMS) based on already existing standards like ISO 27001 [12] or the NIST cyber security framework. Furthermore, technical security requirements are specified distinguishing different security levels for industrial automation and control systems, and also for the used components. The standard has been created to address the specific requirements of industrial automation and control systems.

As shown in Figure 5, different parts of the IEC62443 standard are grouped into four clusters, covering:

– common definitions and metrics;
– requirements on setup of a security organization (ISMS related, comparable to ISO 27001 [12]), as well as solution supplier and service provider processes;
– technical requirements and methodology for security on system-wide level, and
– requirements on the secure development lifecycle of system components, and security requirements to such components at a technical level.

The framework parts address different roles over different phases of the (system) lifecycle: The operator of an automation system operates the automation and control system that has been integrated by the system integrator, using components of product suppliers. In the set of corresponding documents, security requirements are defined, which target the solution operator and the integrator but also the product manufacturer.

According to the methodology described in IEC 62443 part 3-2, a complex automation system is structured into security zones that are connected by and communicate through so-called “conduits” that map for example to the logical network protocol communication between two security zones, see Figure 6.

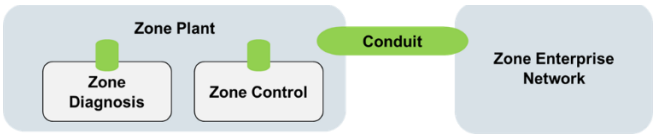

Figure 6. Zones and Conduits

Moreover, this document defines Security Levels (SL) that correlate with the strength of a potential adversary as shown in Figure 7. To achieve a dedicated SL, the defined requirements have to be fulfilled. IEC 62443 part 3-3 defines system security requirements. It supports focusing only on certain facets of security. The security requirements defined by IEC 62443 helps to ensure that all relevant aspects are addressed.

Part 3-3 of IEC 62443 [14], addressing an overall automation system, is in particular relevant for the system integrator. It defines seven foundational requirements that group specific requirements of a certain category:

– FR 1 Identification and authentication control
– FR 2 Use control
– FR 3 System integrity
– FR 4 Data confidentiality
– FR 5 Restricted data flow
– FR 6 Timely response to events
– FR 7 Resource availability

For each of the foundational requirements, there exist several concrete technical security requirements (SR) and requirement enhancements (RE) to address a specific security level. In the context of communication security, these security levels are specifically interesting for the conduits connecting different zones. Related security requirements are defined for the components of an industrial automation and control

system in IEC 62443 part 4-2 [15], addressing in particular component manufacturers. The definition of security requirements distinguishes different categories of components, which are "software application", "embedded device", "host device", and "network device".

Four Security Levels (SL1, SL2, SL3, SL4) are defined that correlate with the strength of a potential attacker as shown in Figure 7. The targeted security level of a zone of the industrial automation and control system is determined based on the identified risk. This allows to tailor the security requirements to the specific needs of an industrial automation and control system.

To reach a dedicated security level, the System Requirements (SR) and potential Requirement Enhancements (RE) defined for that security level have to be fulfilled. The standard foresees that a security requirement can be addressed either directly, or by a compensating countermeasure.

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 7. IEC 62443 defined Security Level [9]

The concept of compensating countermeasures allows to reach a certain security level even if some requirements cannot be implemented directly, e.g., as some components do not support the required technical features. This approach is in particular important for existing industrial automation and control systems, so called "brown-field installations", as existing equipment can be continued to be used.

The security level of a zone or a conduit (a conduit connects zones) is more precisely a security level vector with seven elements. The elements of the vector designate the security level for each foundational requirement. This allows defining the security level specific for each foundational requirement. If, e.g., confidentiality is no security objective within a zone, the security level element corresponding to FR4 "Data confidentiality" can be defined to be SL1 or even none, although SL3 may be required for other foundational requirements (e.g., for FR1, FR2, and FR3). So, the resulting security level vector for a zone could be SL=(3,3,3,1,2,1,3) or SL=(2,2,2,0,1,1,0). The seven elements of the SL-vector correspond to the seven foundational requirements, so that the security level $SL_{FR(i)}$ can be defined separately for each foundational requirement FR(*i*), i.e., SL = (sl_{FR1}, sl_{FR2}, sl_{FR3}, sl_{FR4}, sl_{FR5}, sl_{FR6}, sl_{FR7}).

Different types of SL vectors are distinguished, depending on the purpose:

- SL-T: A target security level vector is defined by the IACS operator based on his risk assessment, defining which security level shall be achieved by each zone and conduit.
- SL-A: The achieved security level vector designates the current status, i.e., the security level that is actually achieved by each zone and conduit. In particular for brown-field installations, it is common that a targeted security level cannot be set-up immediately. The gap between the targeted and the actually achieved security level can be made transparent.
- SL-C: The security level capability describes the reachable security level a component is capable of, if properly configured, without additional compensating counter measures employed. This also means that depending on the SL-T not all security features of a component may be used in certain installations.

C. IEC 62443 Integrity Requirements

One of the seven foundational security requirements defined in Part 3-3 of IEC 62443 [14], targets specifically integrity.

Integrity requirements cover the following areas:

- Overall system integrity
- Communication integrity
- Device integrity

The following examples from IEC 62443-3-3 [14] illustrate some of the integrity-related requirements:

- FR3, SR3.1 Communication integrity: "The control system shall provide the capability to protect the integrity of transmitted information".
- FR3, SR3.4 Software and information integrity: "The control system shall provide the capability to detect, record, report and protect against unauthorized changes to software and information at rest."
- FR3, SR3.8 Session integrity: "The control system shall provide the capability to protect the integrity of sessions. The control system shall reject any usage of invalid session IDs."
- FR5, SR 5.2 Zone boundary protection: "The control system shall provide the capability to monitor and control communications at zone boundaries to enforce the compartmentalization defined in the risk -based zones and conduits model."

D. Practical Application of IEC 62443

The standard IEC 62443[11] is applied successfully by operators, integrators, and manufacturers in various projects. However, it is common that the security documentation and technical designs of real-world deployments are not made public or shared with competitors. Still, some examples for applying the IEC 62443 standards have been made available

publicly: An example of a possible application of the IEC 62443 standard to an Ukrainian power plant gives some insight concerning how the standard can be applied in a concrete setting [16]. In particular, it shows that a sound, comprehensive security concept is needed that covers security requirements broadly and at a consistent level addressing both, organizational/procedural and technical security requirements. The German industrial association "Zentralverband Elektrotechnik- und Elektronikindustrie e.V." (ZVEI) published an overview document on IEC 62443 that includes an example, showing the application to a simplified automation system [17]. A further example is provided by a blueprint for the design of secure substations in the power system domain [25]. This blueprint has been certified as IEC 62443-2-4 and IEC 62443-3-3 compliant [26].

E. Resilience

Being resilient means to be able to withstand or recover quickly from difficult conditions [2]. It shifts the focus of "classical" IT/OT security, which puts the focus on preventing, detecting, and reacting on cyber-security attacks, to the aspect to continue to deliver an intended outcome despite an adverse cyber attack is taking place, and to recover quickly to regular operation. More specifically, resilience of a system is the property to be resistant to a range of threats and withstand the effects of a partial loss of capability, and to recover and resume its provision of service with the minimum reasonable loss of performance [3]. It has been addressed in telecommunications, ensuring that subscribers can continue to be served even when one line is out of service. Bodeau and Graubart [6] define resilience guidelines for providers of critical national telecommunications infrastructure in the UK. Kott and Linkov [7] have compiled a book of different contributions addressing various aspects of cyber resilience in networks and systems. Besides an overview on cyber security, metrics to quantify cyber resilience, approaches to assess, analyze and to enhance cyber resilience are described. The notion of resilience is related to risk management, and also to robustness. Risk management, the "classical" approach to cyber security, identifies threats and determines the risk depending on probability and impact of a potential attack. The objective is to put the focus of defined security measures on the most relevant risks. Resilience, however, puts the focus on a reduction of the impact, so that the system stays operational with a degraded performance or functionality even when it has been attacked successfully, and to recover quickly from a successful attack. Robustness is a further related approach that tries to keep the system operational *without* a reduction of the system performance [7], i.e., to withstand attacks.

Figure 8 illustrates the concept of cyber resilience: Even if an attack is carried out, the impact on the system operation, i.e., the performance or functionality of the system, is limited. The effects of an attack are "absorbed", so that the system stays operational, but with limited performance or functionality. A recovery takes place to bring the system up to the regular operation. In adaptation of resilience, the system might be enhanced to better prepare for future attacks. In a cyber-physical environment, a main objective is that the CPS stays operational and that its integrity is ensured. In the context of an industrial automation and control system, that means that (only) intended actions of the system in the physical world continue to take place even when the automation and control system of the CPS should be attacked.

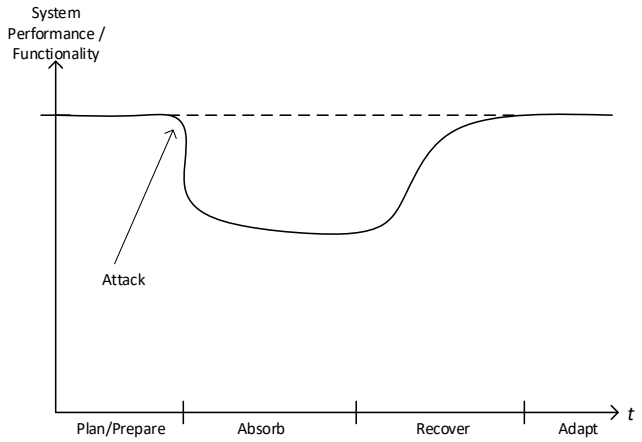

Figure 8. Concept of Cyber Resilience

IV. PROTECTING THE CPS PHYSICAL WORLD INTERFACE

Well-known IT security technologies are encryption and access control, protecting data at rest, in transit, and partly even data in use. In cyber-physical systems, this is not enough. Also, the interface between the OT part (automation systems) and the physical world has to be protected, limiting the potential danger that an automation system can have on the physical world when it is attacked. A successful attack on the automation system or control network can have an impact on the physical world [4].

This section describes the concept of a "physical world firewall" that limits the access to the physical world from OT automation systems. The objective is to increase the resilience of cyber-physical systems, by limiting the impact of an attacked automation system on the physical world. It can be seen as a specific approach for increasing cyber resilience, to design for reversibility. This approach means in general that a cyber physical system should be designed in a manner that allows to revert to a safe mode after components have failed or have been compromised [7]. The approach of a physical world firewall, described in the following section, can be both integrated in automation components, or realized as an add-on component to enhance resilience of existing cyber-physical systems (brown-field). It protects the interface between the control system and the physical world, limiting the possible impact of a successful attack on the physical world.

A. Physical-World Firewall

The main idea or the approach is to filter the communication between sensors and actuators on one side, and the control equipment on the other side. This can be called physical-world firewall. It limits in which way a control system, potentially under attack, can impact a physical system in the real world. The filtering takes place directly at the input/output interface, so that it is independent from the

software-based functionality of the automation component. Conceptually, it can be considered as a physical world reference monitor to control access to the physical world based on a defined access control policy [8]. However, the physical world firewall described here would be realized independent of the software-based control functionality to ensure that is effective even if the software would be manipulated.

Similar as a communication firewall for data traffic that analyzes and filters data packets (IP packets and IP-based communication, filtering based on network addresses and used protocols), here the actuator and sensor signals are filtered, so that only signals allowed by the signal filter policy are provided.

The allowed signal ranges and dynamic parameters are monitored and limited. If the signal filtering policy is violated, the signal cannot be simply dropped like an IP packet. Instead, a replacement signal is provided. The replacement signal may be a fixed default value, or a clipped maximum/minimum value that is within the allowed value range, or it may be an out-of-range signal or a high-impedance signal that will be detected by an actuator as failure signal, so that the actor can react accordingly).

Figure 9 shows an automation component with an integrated Cyber Physical Controlled Input / Output Interface (CPC IO) that realizes a physical-world firewall functionality. Each input/output channel is monitored separately by the "Value Check" component: It verifies whether the current sensor input value or the current actuator output values are within the given allowed range, and thus are compliant with the defined filtering policy *Pol*.

Besides the value range, also further parameters can be calculated and checked against the defined filtering Policy *Pol*, e.g., statistical parameters as average and variation, and dynamic parameters as a first order or second order derivation, or a transformation as a Fourier transformation. Besides the actual input/output signals, also further data relating to the current operating state of the CPS can be used.

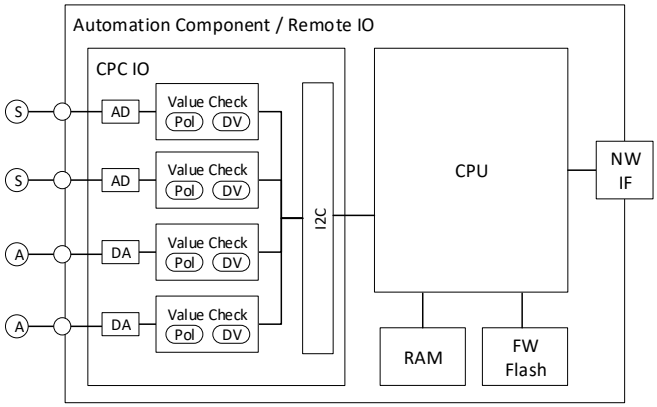

Figure 9. Automation Component with Integrated Physical World Firewall

If the policy is met, the value is allowed, i.e., the unmodified signal is forwarded. Otherwise, the configured default value (DV) is provided as replacement signal, ensuring the CPS system stays operational. It is possible to lock the input/output interface in the case of a detected policy violation. The lock may be permanent, but preferably it can be reset at a reboot of by a manual user interaction.

It is possible that the CPU performs an integrity check as part of a secure boot process or during operation. The CPU subsystem can authenticate towards its CPC IO block after a successful self-integrity check. The CPC IO block can configure a policy depending on the integrity check status of the CPU, limiting the access to the physical world for a manipulated CPU subsystem.

A variant is shown in Figure 10, where the signals of multiple input/output channels are checked in combination. This allows to perform cross-checks between sensor and actuator signals. Moreover, if this approach is applied in a distributed system, it allows to take certain properties of potentially different sensors/actuators into account.

Specifically, if the sensors/actuators used are a mixture of standard (legacy) and specifically hardened, trusted sensors, a potential security assertion can be used in the evaluation of the signals, giving the trusted sensor a higher weight in the evaluation. This is especially advantageous if a larger number of legacy sensors/actuators is already deployed and secure siblings are installed as add-on in a stepwise manner. More information on the basic concept is described in [9].

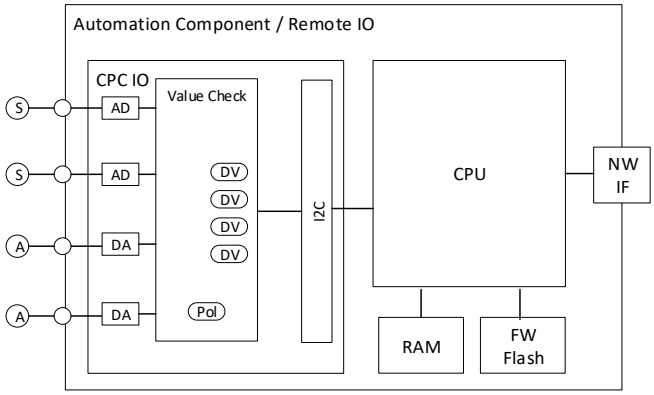

Figure 10. Automation Component with Integrated Physical World Firewall

Both Figure 9 and Figure 10 showed the physical world firewall as an integrated functionality of an automation component. However, it is possible as well to realize the physical world firewall as an add-on component to an existing automation component. This add-on component monitors input/output signals of the automation component between the automation component and the actual sensor/actuator connected to the automation component. The signal is replaced with a replacement signal if the currently observed signal is not complaint with the defined policy *Pol*.

A physical-world firewall realized as add-on component to already existing and deployed automation components can be used in particular within brownfield CPS. A stepwise migration of existing brownfield CPS towards systems with a higher resilience under attack is supported, as the already deployed components of the CPS have not to be replaced.

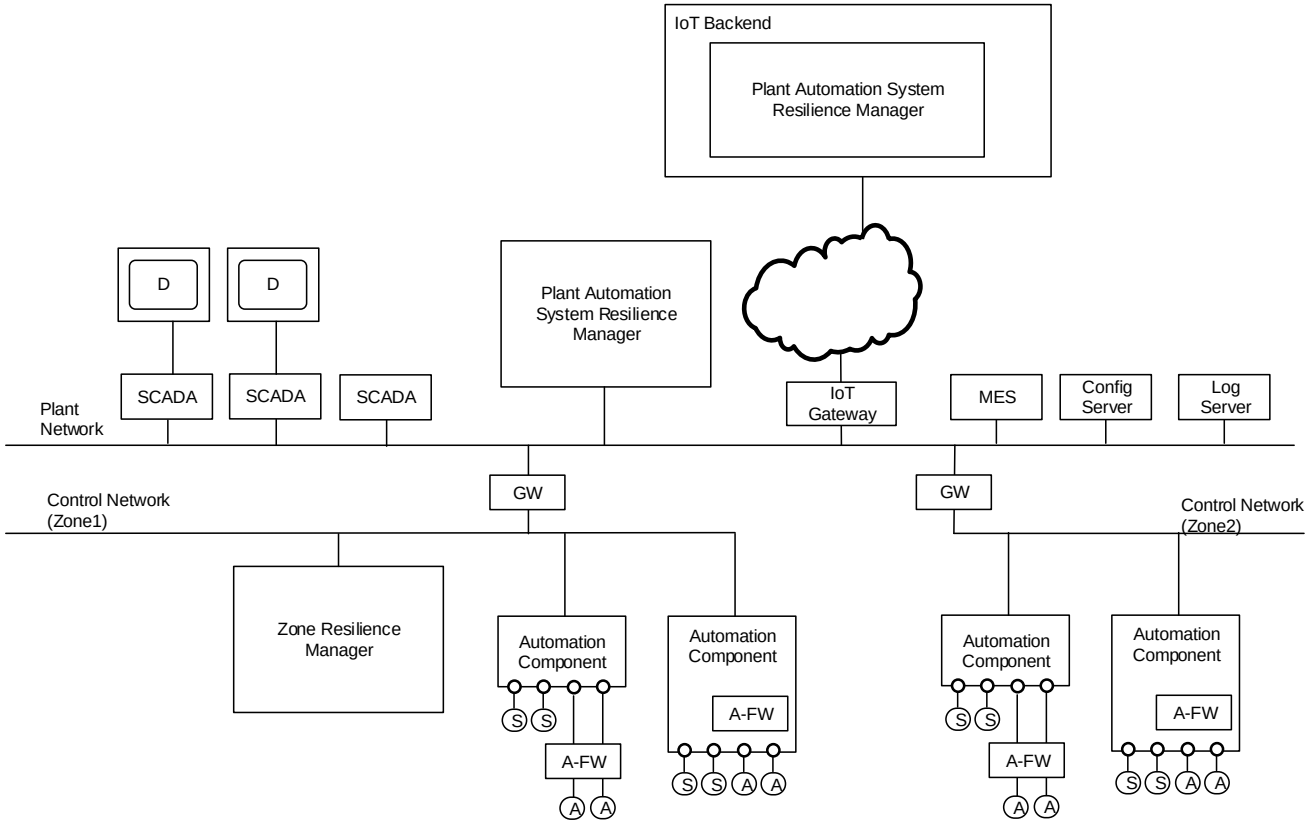

Figure 11. Dynamic Resilience Management

B. *Dynamic Resilience Management*

The policy of the physical-world firewall can be adapted dynamically, depending on the current operating state of the CPS. This allows to restrict the possibility to influence the physical world even more strictly, as the current state of the production system and the currently performed production step, e.g., cooling or filtering a fluid, can be reflected in the current configuration of the physical-world firewalls.

Resilience managers determine the physical-world firewall policy dynamically, depending on the current state and context of the CPS, see Figure 11.

Resilience managers adapt during operation the current policy configuration of the physical-world firewalls.

The policy adaptation performed by resilience managers can use in particular the following information:

- The current state of the physical world, as obtained by trusted sensor nodes [9].
- The current production batch, the current production step, operating state (e.g., standby, preparation, active, service, alarm). In real-world deployments, the information may be obtained from a Manufacturing Execution System (MES).
- Cyber attacks detected by an integrity monitoring system or an intrusion detection system, supervising the CPS.

The dynamic adaptation allows to enforce tight physical-world firewall policies depending on the current system state and operation.

C. *Policy Adaptation for Dynamically Reconfigurable CPS*

Cyber-physical systems and industrial automation systems are often rather static. After being put into operation, changes to the configuration happen only rarely, e.g., to replace a defect component, or to install smaller upgrades during a planned maintenance window. To cope with increasing demands for flexible production and increased productivity, also CPS will become more dynamic, allowing for reconfiguration during regular operation. Such scenarios for adaptable, reconfigurable production have been described in the context of Industry 4.0 [19].

An integrity monitoring system for a reconfigurable CPS has to be adapted to the current configuration. Similar as for dynamic resilience management, the policies for physical world firewalls can be adapted depending on the current CPS configuration. The information of the current configuration is usually managed already as part of collecting, storing, and validating production data that describes the production

process, so that the information for adapting the policies is available already.

D. Physical World Integrity Monitoring

A further source of information for adapting the filter policy is monitoring the automation site for physical security, using alarm systems, e.g., physical access control and closed-circuit television (CCTV) cameras. If the alarm system detects some unexpected situation, e.g., an intruder, the filter policies of the physical world firewalls can be reconfigured to limit possible damages.

Furthermore, the physical operational properties of automation machinery (e.g., drives, pumps) can be monitored. The acoustic emissions (vibrations) of machines as well as power consumption profile (power fingerprinting) can be monitored. Signal processing algorithms including machine learning (artificial intelligence) can be used to determine whether the machinery is in a normal or exceptional operational state. Also, specific actuations in the physical world can be performed that encode integrity measurements of software and data in control operations, realizing integrity attestation by physical-world actuation signals [18].

If it is detected that the machinery is operating outside the expected operational boundaries, the filter policy of the physical world firewalls can be adapted to limit the impact of the automation system on the physical world accordingly. A restricted physical-world filter policy can be configured dynamically that is foreseen for detected integrity violations.

E. Authenticating Physical Signals

In data communication, the sender of a data packet can be identified by an identifier, e.g., an internet protocol (IP) address or a media access control (MAC) address. The sender may be authenticated cryptographically. A data firewall can filter data packets depending on address information and content. In the physical world, the source of a signal can in general not be identified by an explicit identifier included in the data communication. However, the source is implicitly based on the cabling.

A higher level of confidence can be achieved by performing signal authentication. The sender of a signal can be identified by a sender-specific fingerprint information, e.g., a noise signal. Furthermore, it is possible to actively add a signal marker (signal watermarking), e.g., a coded spread-spectrum signal [20][21][22]. This allows to identify the source of a signal by evaluation properties of the signal. The physical world firewall can identify signals not having the expected fingerprint and block them, i.e., substitute them with a replacement signal. A (physical) signal cannot simply be blocked by not forwarding it. The replacement signal may be a regular signal value, or a specific out-of-range signal value.

The coded spread spectrum signal (signal watermark) can be added to the actual measurement signal close to the analog sensor by adding the watermarking noise signal. However, it is also possible to add actuators in the physical world of the CPS that imprint a watermarking noise signal in the physical world, e.g., by mechanical actuator. Thereby, already deployed sensors (brownfield installation) can capture the watermarking signal, and the sensor measurements can be verified. While having some similarities to the approach described by Ghaeini, Chan, et al. [18], here, the physical world watermarking ensures the reliable identification/authentication of physical signals.

F. Defining Policies for Physical World Firewalls

Even for conventional firewalls filtering network communication, the definition and testing of firewall policies is a huge challenge. The level of security of a network that is actually achieved depends heavily on the ability to manage the available security mechanisms effectively and consistently [23]. This is the case in particular when several firewalls are deployed that have to be configured consistently, and when involving multiple administrative domains. The administration has to be practical, i.e., both efficient and effective, also in such complex environments, with frequent changes and with the complexity of networks consisting of thousands of users and components.

The same applies to policies for physical world firewalls. A further specific side condition is that properties of the physical world have to be understood to come-up with an appropriate policy. This requires a good understanding of the automation system and the physics of the controlled system. A manual configuration of such policies will hardly be practical in real-world deployments. As with other security mechanisms, also physical world firewalls will be introduced stepwise, starting with less critical parts of the CPS and with simple policies to avoid unexpected negative impacts on the regular operation.

For the practical definition of the policies, two approaches seem promising:

- CPS simulation: One important aspect of Industry 4.0 is a digital twin of the physical system that allows to perform simulations in the digital world. Here, the CPS can be simulated under all foreseen operational conditions to derive the filter policies permitting all signals that can be expected in foreseen operational conditions. Also, specific attack scenarios can be simulated.
- Machine learning: The policies can be learnt, similar to network firewalls, where during a learning phase, the policy is automatically determined.

These approaches for determining the filter policies automatically can be enhanced with hand-crafted, manually defined filter policies for interfaces to highly critical physical world components, or for highly critical automation steps. Those tight policies can be adapted specifically to the purpose and foreseen usage of the automation component. So, tight physical world firewall policies can be defined based on a risk assessment, protecting the most relevant components and automation steps.

V. EVALUATION

The security of a cyber system can be evaluated in practice in various approaches and stages of the system's lifecycle:

- Threat and Risk Analysis (TRA, also abbreviated as TARA) of a cyber physical system (for a system being

under design or in operation). In a TRA, possible attacks (threats) on the system are identified. The impact that would be caused by a successful attack (threat) and the probability that the attack happens are evaluated to determine the risk of the identified threats. The risk evaluation allows to prioritize the threats, focusing on the highest, most relevant risks and to define corresponding security measures. Besides technical measures, also organizational and personal security measures can be defined.

- Security checks can be performed during operation or during maintenance windows to determine key performance indicators (e.g., check compliance of device configurations). It can be verified that the defined security measures are in fact in place, and areas requiring increased attention can be identified.
- Security testing (penetration testing, also called pentesting for short) can be performed for a system that has been built, but that is currently not in operation. A pentest can usually not be performed on an operational automation and control system, as the pentest could endanger the reliable operation of the system. Pentesting can be performed during a maintenance window when the physical system is in a safe state, or using a separate test system. The non-operational system is attacked by "white hat" hackers to identify vulnerabilities that need to be addressed.
- Security testing can be performed also on a digital representation of a target system, e.g., a simulation in the easiest case. This digital representation is also called "digital twin". This allows to perform security checks and pentesting for systems that are not existing yet physically (design phase), or to perform pentesting of operational systems in the digital world without the risk of disturbing the regular operation of the real-world system.

A holistic protection concept has to address measures for all three discussed phases: protect, detect, and react. No single measure or security technology alone can result in an adequate security level. It is always a set of measures that, when used in combination, can reduce the overall risk to an acceptable level.

The security measures presented in this paper, acting on the interface between the cyber world and the physical world, provide an additional security measure that can be used as part of a defense-in-depth security concept. The protection is complementary to well-known security measures that focus on the IT/cyber part, as it operates directly at the interface towards the physical world, not on computer-based control functions as conventional IT security technologies. Even if all security measures in the pure IT/cyber world fail, still the impact on the physical world can be controlled. It can serve as "last line of defense", allowing to connect cyber systems from the physical world in a tightly controlled way.

A limitation for all evaluations of the effectiveness of an overall security architectural design and of individual security measures is the fact that the threat landscape of attacks seen in practice continuously evolves. Therefore, it is required that a security design allows for being updated to address new attacks. This aspect is in particular important for CPS and automation and control systems having typically a long life-time of typically 10 to 30 years. The defined concept of physical world firewalls supports an update not only to already existing brownfield installations, but also to enhance the security robustness of long-lived systems during operation without directly affecting the control functionality. As CPS are often subject to regulatory approvals, having security measures that can be updated and enhanced along the lifetime without directly affecting regulatory approvals of the control functions is advantageous.

As long as the proposed technology has not been proven in a real-world operational setting, it can be evaluated conceptually by analyzing the impact that the additional security measure would have on the identified residual risks as determined by the TRA of the CPS. A TRA identifies threats against a system, and determines the risk depending on probability and impact. The general effect of the presented security measure is that the impact of a threat, i.e., a successful attack, on the physical world controlled by the CPS is reduced. Whatever attack is ongoing on the IT-based automation and control system, still the possible impact on the real, physical world is limited. So, the measure helps to reduce the risk of threats having an impact on the physical world.

However, TRAs for real-world CPS are not available publicly. Nevertheless, an illustrative example may be given by a chemical production plant performing a specific process like refinery, or a factory producing glue or cement. If the plant is attacked, the attack may target to destroy the production equipment by immediately stopping the process leading to physical hardening of the chemicals / consumables and thus to a permanent unavailability of the production equipment. In this case, trusted sensors could be used to detect a falsified sensor signal, and the physical-world firewall can be used to limit actions in the physical world. Thereby, a physical damage of the production equipment can be avoided. If needed, a controlled shutdown of the production site can be performed.

As the evaluation in a real-world CPS requires significant effort, and as attack scenarios cannot be tested that could really have a (severe) impact on the physical world, a simulation-based approach or using specific test-beds are possible approaches, allowing to simulate or evaluate in a protected test-bed the effect on the physical world of certain attack scenarios with compromised components. The simulation would have to include not only the IT-based control function, but also the physical world impact of an attack. Using physical-world simulation and test beds to evaluate the impact of attacks have been described by Urbina, Giraldo et al. [24].

A major advantage of the physical-world firewall is the property that it can be added to existing brownfield deployments. Legacy equipment, may be 10 or even 20 years in the field, not even been designed with security in mind, and without getting patches. In such cases, the physical-world firewall can be used as an "add-on" security measure for an

existing CPS. It can be used as compensating countermeasure to address security requirements defined by industrial security standards like IEC 62443-3.3 [14], where conventional cyber security measures cannot be deployed. However, it can be used also as additional layer of defense in CPS having state-of-the-art security measures integrated, thereby increasing the level of protection even further. The conceptual advantage that the protection acts on a different layer than conventional IT security mechanisms provides an additional, independent layer of defense. As for all security technologies, the confirmation for the actual effectiveness has to come from tests and experience real-world application, starting with smaller pilot tests in real deployments.

VI. Conclusion

With ubiquitous machine-oriented communication, e.g., the Internet of Things and interconnected cyber physical systems (CPS), the integrity of the operation of technical systems is becoming an increasingly important security objective. Protecting such systems against intentional attacks to ensure a reliable operation is demanded by operators, as well as by regulation. There is a need for enhanced protection that can be applied practically both to already deployed installations, where often IT-based functionality cannot be updated practically, as well as to new CPS, which are increasingly open and dynamic.

A CPS comprises the operational cyber-technology and the physical world with which the system interacts. Both parts have to be covered by a security concept and solution. Cyber security puts the focus traditionally on the cyber-part, i.e., on the IT-based automation and control systems. The security of the physical part, like machinery, is protected often by physical and organizational security measures, only. This is challenging for dynamically changing cyber physical systems, that come with the Industrial Internet of Things (IIoT) and Industry 4.0. Cyber systems will become more and more open and dynamic to support flexible production down to lot size 1 (plug-and-work reconfiguration of manufacturing equipment), and to support a flexible adaptation to changing needs like market demand and personalized products.

This paper presented a concept for a new approach that enhances the achieved level of security by protecting the interface between the IT-based cyber-part and the physical world, thereby enhancing the resilience of a CPS being under attack. The CPS may even continue to operate under attack, as the possible negative impact on the physical world is restricted. This allows also to ensure a high availability of the automation system, even under attack, as the automation system has not to be shut down.

The proposed new layer of protection can be applied to new installations (greenfield), e.g., to address the risk of installing malware during update of the software-based functionality. More importantly, it can as well be applied as add-on to already deployed installations (brownfield). It realizes an additional, independent level of protection that can be deployed and updated independently of the actual control systems of the legacy system. Therefore, it can also be applied when a legacy IT-based control system of a CPS cannot be updated with current cyber security technology. This is a demanding problem in many installed CPS, as they are often in use for several decades and are subject to regulations that make updates complicated or even impossible. The proposed solution can be introduced in a complex CPS in a stepwise way, starting with most critical physical world interfaces. Also, the filtering policies can be coarse in an initial usage phase, and it can be updated with increasing sophistication depending on observed attacks, and reflecting the intended operation of the specific CPS and its current operation mode.

References

[1] R. Falk and S. Fries, “Enhancing Resilience by Protecting the Physical-World Interface of Cyber-Physical Systems”, The Fourth International Conference on Cyber-Technologies and Cyber-Systems CYBER 2019, September 22, 2019 to September 26, 2019 - Porto, Portugal, [Online]. Available from: https://www.thinkmind.org/index.php?view=article&articleid=cyber_2019_1_20_80033 2020.05.13

[2] P. England, R. Aigner, A. Marochko, D. Mattoon, R. Spiger, and S. Thom, “Cyber resilient platforms”, Microsoft Technical Report MSR-TR-2017-40, Sep. 2017, [Online]. Available from: https://www.microsoft.com/en-us/research/publication/cyber-resilient-platforms-overview/ 2020.05.13

[3] Electronic Communications Resilience&Response Group, “EC-RRG resilience guidelines for providers of critical national telecommunications infrastructure”, version 0.7, March 2008, available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/62281/telecoms-ecrrg-resilience-guidelines.pdf 2020.05.13

[4] D. Urbina, J. Giraldo, N. O. Tippenhauer, and A. Cardenas, “Attacking fieldbus communications in ICS: applications to the SWaT testbed”, Singapore Cyber-Security Conference (SG-CRC), IOS press, pp. 75–89, 2016, [Online]. Available from: http://ebooks.iospress.nl/volumearticle/42054 2020.05.13

[5] C. C. Davidson, T. R. Andel, M. Yampolskiy, J. T. McDonald, W. B. Glisson, and T. Thomas, “On SCADA PLC and fieldbus cyber security”, 13th International Conference on Cyber Warfare and Security, National Defense University, Washington, DC, pp. 140–148, 2018

[6] D. Bodeau and R. Graubart, “Cyber resiliency design principles”, MITRE Technical Report, January 2017, [Online]. Available from: https://www.mitre.org/sites/default/files/publications/PR%2017-0103%20Cyber%20Resiliency%20Design%20Principles%20MTR17001.pdf 2020.05.13

[7] A. Kott and I. Linkov (Eds.), “Cyber Resilience of Systems and Networks”, Springer, 2019

[8] E. B. Fernandez, M. VanHilst, D. laRed Martinez, and S. Mujica, An Extended Reference Monitor for Security and Safety, 5th Ibero-American Congress on Information Security, Montevideo, Uruguay, November 2009, [Online]. Available from: http://www.criptored.upm.es/cibsi/cibsi2009/docs/Papers/CIBSI-Dia2-Sesion4(3).pdf 2020.05.13

[9] R. Falk and S. Fries, “Enhancing integrity protection for industrial cyber physical systems”, The Second International Conference on Cyber-Technologies and Cyber-Systems, CYBER 2017, pp. 35–40, November 12 - 16, 2017, Barcelona, Spain, [Online]. Available from: http://www.thinkmind.org/index.php?view=article&articleid=cyber_2017_3_30_80031 2020.05.13

[10] European Commission, "The directive on security of network and information systems (NIS Directive)", [Online]. Available from: https://ec.europa.eu/digital-single-market/en/network-and-information-security-nis-directive 2020.05.13

[11] IEC 62443, "Industrial automation and control system security" (formerly ISA99), [Online]. Available from: http://isa99.isa.org/Documents/Forms/AllItems.aspx 2020.05.13

[12] ISO/IEC 27001, "Information technology – security techniques – Information security management systems – requirements", October 2013, [Online]. Available from: https://www.iso.org/standard/54534.html 2020.05.13

[13] NIST, "Framework for Improving Critical Infrastructure Cybersecurity", Version 1.1, April 16, 2018, [Online]. Available from: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf 2020.05.13

[14] IEC 62443-3-3:2013, "Industrial communication networks – network and system security – Part 3-3: System security requirements and security levels", Edition 1.0, August 2013

[15] IEC 62443-4.2, "Industrial communication networks - security for industrial automation and control systems - Part 4-2: technical security requirements for IACS components", CDV:2017-05, May 2017

[16] P. Bock, J.-P. Hauet, R. Françoise, and R. Foley, "Ukrainian power grids cyberattack - A forensic analysis based on ISA/IEC 62443", ISA InTech magazine, 2017, [Online]. Available from: https://www.isa.org/templates/news-detail.aspx?id=152995 2020.05.13

[17] ZVEI, "Orientation guideline for manufacturers on IEC 62443", "Orientierungsleitfaden für Hersteller zur IEC 62443" [German], ZVEI Whitepaper, 2017, [Online]. Available from: https://www.zvei.org/presse-medien/publikationen/orientierungsleitfaden-fuer-hersteller-zur-iec-62443/ 2020.05.13

[18] H.R. Ghaeini, M. Chan, R. Bahmani, F. Brasser, L. Garcia, J. Zhou, A.-R. Sadeghi, N. O. Tippenhauer, and S. Zonouz, "PAtt: Physics-based Attestation of Control Systems", 22nd International Symposium on Research in Attacks, Intrusions and Defenses, USENIX, September 23-25, 2019, [Online]. Available from: https://www.usenix.org/system/files/raid2019-ghaeini.pdf 2020.05.13

[19] Plattform Industrie 4.0, "Industrie 4.0 Plug-and-produce for adaptable factories: example use case definition, models, and implementation", Plattform Industrie 4.0 working paper, June 2017, [Online]. Available from: https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Industrie-40-Plug-and-Produce.pdf , 2020.05.13

[20] T. Hupperich, H. Hosseini, and T. Holz, "Leveraging sensor fingerprinting for mobile device authentication", International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, LNCS 9721, Springer, pp. 377–396, 2016, [Online]. Available from: https://www.syssec.ruhr-uni-bochum.de/media/emma/veroeffentlichungen/2016/09/28/paper.pdf 2020.05.13

[21] H. Bojinov, D. Boneh, Y. Michalevsky, and G. Nakibly, "Mobile device identification via sensor fingerprinting", arXiv:1408.1416, 2016, [Online]. Available from: https://arxiv.org/abs/1408.1416 2020.05.13

[22] P. Hao, "Wireless device authentication techniques using physical-layer device fingerprint", PhD thesis, University of Western Ontario, Electronic Thesis and Dissertation Repository, 3440, 2015, [Online]. Available from: https://ir.lib.uwo.ca/etd/3440 2020.05.13

[23] R. Falk and M. Trommer, "Integrated Management of Network and Host Based Security Mechanisms," 3rd Australasian Conference on Information Security and Privacy, ACISP98, pp. 36-47, July 13-15, 1998, LNCS 1438, Springer, 1998

[24] D. Urbina, J. Giraldo, A. Cardenas, N. O. Tippenhauer, J. Valente, M. Faisal, J. Ruths, R. Candell, and H. Sandberg, "Limiting The Impact of Stealthy Attacks on Industrial Control Systems," ACM Conference on Computer and Communications Security (CCS), Vienna, Austria, 2016

[25] Siemens, "Secure Substation Manual", [Online]. Available from: https://www.siemens.com/download?DLA20_114 2020.05.13

[26] Siemens, "Digital Substation Cyber Security", [Online]. Available from: https://www.siemens.com/download?DLA13_3680 2020.05.13