CHEMICAL ENGINEERING TRANSACTIONS
VOL. 111, 2024
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: Valerio Cozzani, Bruno Fabiano, Genserik Reniers

ISBN 979-12-81206-11-3; **ISSN** 2283-9216
DOI: 10.3303/CET24111031

Critical Cybersecurity Scenarios in Drinking Water Treatment Plants

Matteo Iaiani*, Alessandro Tugnoli, Valerio Cozzani

LISES – Department of Civil, Chemical, Environmental, and Materials Engineering, Alma Mater Studiorum – University of Bologna, via Terracini n.28, 40131 Bologna, Italy
matteo.iaiani@unibo.it

The increasing interconnectivity with external networks and the higher reliance on digital systems make chemical and process industries, including waste and drinking water treatment plants, more vulnerable to cyber-attacks. Historical evidence shows that these attacks have the potential to cause events with severe consequences on property, people, and the surrounding environment, posing a serious threat. While the risks deriving from the malicious manipulation of the Basic Process Control System (BPCS) and the Safety Instrumented System (SIS) in chemical and Oil&Gas facilities have been systematically analysed in the available literature, including previous works of the Authors, the analysis of the consequences of cyber-attacks to drinking water treatment plants has not been conducted to date. To fill this gap, in the present study the methodology POROS 2.0 (Process Operability Analysis of Remote manipulations through the cOntrol System) developed by the Authors was applied to a drinking water treatment plant, providing valuable insights on possible critical scenarios originated by cyber-attacks in these facilities.

1. Introduction

In recent years, the field of industrial cybersecurity has gained increasing attention due to the growing digitization and connectivity of critical infrastructure facilities, including chemical, process, and water treatment plants, which play a pivotal role in ensuring the safety and well-being of society, and as such, they are entrusted with the responsibility of managing complex processes and systems (CCPS, 2022). The integration of digital technologies and the adoption of Industrial Internet of Things (IIoT) devices have streamlined operations, improved efficiency, and provided better control over these processes. However, this increased reliance on digitalization has inadvertently exposed these plants to a plethora of cybersecurity threats (Iaiani et al., 2020). The cyber-induced explosion of the BTC pipeline in Turkey in 2008 and the ransomware attack on Colonial Pipeline in USA in 2021 are noteworthy examples (Iaiani et al., 2023b).
In this panorama, the ISA/IEC 62443 series of standards provide a systematic and practical approach to evaluate the actual level of cyber risk and to implement proper cybersecurity countermeasures in industrial critical infrastructures. This requires the identification of all the impacts that can result from deliberate malicious attacks to the BPCS (Basic Process Control System) and SIS (Safety Instrumented System), including those on the physical plant (process equipment, storage equipment, interconnections), the evaluation of their consequences, and of their likelihood. However, neither specific methods nor guidelines are provided in the standards to conduct the proposed approach. Similarly, also the classical methodologies dedicated to process plant Security Vulnerability/Risk Assessment (SVA/SRA) such as the VAM-CF, the CCPS SVA, the API RP 780 SRA, and the majority of academic contributions, lack in reproducibility in application or do not assess the actual link between manipulations, consequences, and likelihood (Bajpai & Gupta, 2018).While there are studies in the literature (including previous works of the Authors) focused on the analysis of the possible scenarios that can be triggered by the deliberate manipulation of the BPCS and the SIS in plants where large quantities of dangerous substances are processed and/or stored (Cherdantseva et al., 2016), there are no works with scope on drinking water treatment plants. However, cyber-attacks to this type of industrial infrastructures are deemed

Paper Received: 9 March 2024; Revised: 27 April 2024; Accepted: 6 July 2024
Please cite this article as: Iaiani M., Tugnoli A., Cozzani V., 2024, Critical Cybersecurity Scenarios in Drinking Water Treatment Plants, Chemical Engineering Transactions, 111, 181-186 DOI:10.3303/CET24111031

to be able to generate events with severe consequences on the population and company reputation due to their important role within society.
In this context, the present study is aimed at the identification of critical cybersecurity scenarios in drinking water treatment plants by the application of the cyber-risk identification methodology POROS 2.0 (Process Operability analysis of Remote manipulations through the cOntrol System) developed by the Authors in a previous study (Iaiani et al., 2023b). The results support application of cybersecurity risk assessment in the context of SVA/SRA methodologies and ISA/IEC 62443 series of standards in drinking water treatment plants.

2. POROS 2.0 methodology

POROS 2.0 is a rigorous 8-step procedure (see flowchart in Figure 1) which provides the following outputs: 1) list of security events (SE, such as loss of containment, production outage, etc.) on the physical process system that can be caused by manipulation of BPCS and SIS; 2) severity vector associated with each SE associated to damage to people, asset, environment, and reputation; 3) sets of BPCS and SIS components that need to be manipulated in order to trigger each SE, how they shall be manipulated, and the physical consequences on remotely manipulable components (RMC) in the physical process system; 4) list of the Active/Procedural safeguards (APS) in place, that may prevent/mitigate the attack; 5) list of the Inherent/Passive safeguards (IPS) in place that may prevent/mitigate the attack; 6) credibility score of each attack, based on required plant knowledge level and cyber complexity of the manipulations required to carry it out. In the following, each step is briefly detailed: the reader is referred to Iaiani et al. (2023b) where guiding tables are provided in support of each step.

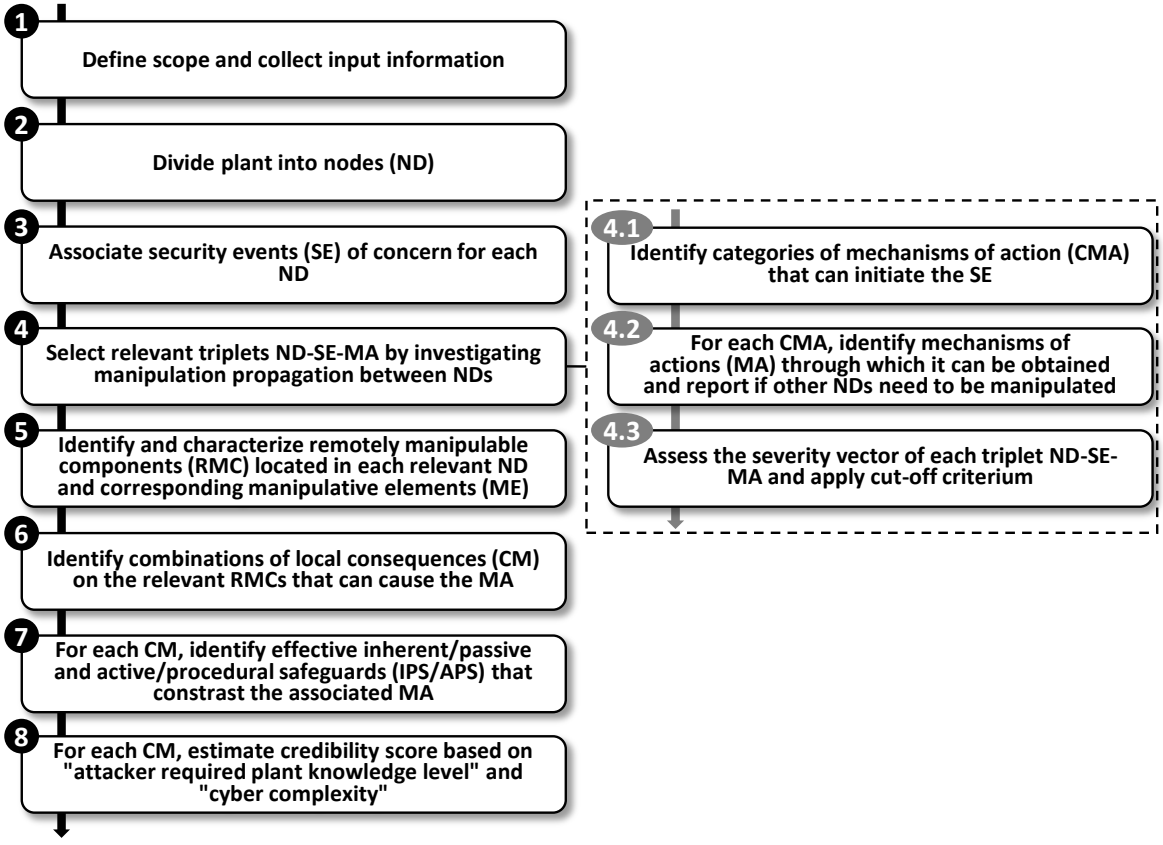

Figure 1: Flowchart of POROS 2.0 methodology (adapted from Iaiani et al. (2023b)*)*

In Step 1 of POROS 2.0 the scope of the analysis is defined (on major accident scenarios only or including also operability issues). The possibility of adopting cut-off criteria to focus the analysis on a limited number of cybersecurity scenarios of concern shall be defined at this step (e.g., worst-case scenarios, or most credible scenarios). Once the scope is defined, the input information is collected (e.g., PFD, P&ID, equipment operating and design conditions, BPCS and SIS logics). In Step 2 of POROS 2.0 the plant is divided into nodes (ND). Only nodes where hazardous materials are handled are of concern in case the scope of the analysis is limited to major accident scenarios, otherwise also the utility nodes are considered.

In step 3 of POROS 2.0 the compatible security events (SE) are identified for each selected ND. A SE is intended as an undesired event that affects the operability and/or the physical integrity of the physical process system under assessment and they shall be selected according to the scope of the analysis (e.g., loss of containment in case of scope on major events, also production outages in case of scope on operability). The reader is referred to the reference source of the methodology (Iaiani et al., 2023b) where a list of possible SEs is provided (from major accident scenarios to scenarios leading to production outage such as “product out of specification” or “activation of ESD/PSD/LSD logic”).

In Step 4 of POROS 2.0 the propagation between NDs of the physical effects of manipulations is investigated and the scenarios relevant for security risk assessment are selected. This step allows to analyze the interdependencies between nodes (as in HazOp application), enabling to catch potential escalation of the consequences of the actions carried out by an attacker in a node. The relevant scenarios are defined in terms of the triplet node - security event - mechanism of action (ND-SE-MA). To guide the systematic application, this step has been divided in 4 sub-steps (steps 4.1 - 4.3), which shall be carried out for each ND. Step 4.1 consists in the identification of the categories of mechanisms of action (CMA) for each SE associated to the ND under assessment. CMAs are general mechanisms that can initiate a SE (e.g., “damage of the construction material of the containment system” is a possible generic CMA that initiates a LOC). Step 4.2 consists in the identification of the specific mechanisms of action (MA) through which each CMA can be obtained in the plant analysed and the ND where such MA shall be carried out. MAs are specific mechanisms (based on the features of the plant analysed) that can initiate a SE (e.g., “inducing high pressure in the separator” is a possible MA which can be grouped into the CMA mentioned above initiating a LOC). It's crucial to emphasize that to obtain a CMA, it might be necessary to perform MAs in nodes other than the one being evaluated (e.g., more flowrate in a node causes high level in an equipment unit in a nearby node): in such situations, information is transferred from one node to another in a manner akin to how deviations spread through various nodes in a conventional HazOp analysis, highlighting interdependencies between nodes. Step 4.3 consists in the assessment of the severity vector associated with each identified triplet ND_i-SE_j-MA_k in the node under assessment. The severity vector evaluates the impact of the SE_j initiated by MA_k in ND_i using four severity levels, i.e., minor (1), medium (2), major (3), and extensive (4), and four targets, i.e., people, assets, environment, and reputation. It is important to underline that the severity vector for each triplet ND-SE-MA shall be estimated considering the effect that such SE initiated through such MA has in the entire plant, not only in the ND where it occurs. Step 4.3 provides also for the selection of the relevant triplets ND-SE-MA. A straightforward cut-off criterium that prioritizes SEs with the most severe impact (i.e., a severity level of 3 or higher for at least one target) is suggested; however, alternative cut-off criteria may be appropriate.

The following steps are then limited in the scope of the selected relevant triplets ND-SE-MA.

In Step 5 of POROS 2.0, the remotely manipulable components (RMC) located in the relevant select NDs and their manipulative elements (ME) are identified and characterized. RMCs are the physical objects in the plant whose operation is regulated by the BPCS and the SIS (e.g., automatic control and shutoff valves, pumps, compressors, etc., and thus every component which is not manual), while MEs are the elements of the BPCS and the SIS that regulate behaviour of RMCs (e.g., PID and PLC controllers and their logics which act on e.g. on valves). Characterization of MEs consists in the definition of the remote manipulations (RMs) that an attacker can carry out on them (e.g., setpoint change on a PID controller). Analogously, characterization of RMCs consists in the definition of the physical changes that occur on them as a consequence of RMs on the ME by which the RMCs are regulated, named local consequences (LC, e.g., the closing of a valve as a consequence of changing the set point of the controller acting on the valve).

In Step 6 of POROS 2.0 the combinations (CM) of local consequences on the RMCs located in the ND that are required to perform the MA are identified.

In Step 7 of POROS 2.0 the Active/Procedural safeguards (APS, e.g., ESD/PSD/LSD logics) and the Inherent/Passive safeguards (IPS, e.g., pressure safety valves) that can be effective in contrasting the MA are identified. Therefore, the manipulations required by the CM and the deactivation of the APSs contrasting the MA to which the CM refers, constitute a “CM+APS attack action” for the triplet ND-SE-MA considered. The CM+APS attack action thus results to be the complete set of all the actions that an attacker has to carry out in order to trigger a specific security event in a node through a specific mechanism of action.

In Step 8 of POROS 2.0 the credibility score of each CM+APS attack action associated to the triplet ND-SE-MA under assessment is evaluated. The score is estimated combining in a matrix a score on two dimensions: the “plant knowledge level” required by the attacker and the “cyber complexity” of the CM+APS attack action. The "plant knowledge level" refers to the level of technical knowledge on the plant under assessment or on similar plants that is required by an attacker to carry out a specific CM+APS attack action, while the "cyber-complexity" of a CM+APS attack refers to how complex the attack is in terms of the number of RMCs that need to be manipulated, the number of zones that need to be accessed in the OT system, and whether a specific sequence

and timing is required. The ranks considered for both “plant knowledge level” and “cyber complexity” and their rationale are proposed by the Authors and provided in Iaiani et al. (2023a).

3. Case study

3.1 Description

The water treatment plant considered in the case study aims to convert wastewater into a primary service (drinking water) through a multi-step process: the block diagram of main process and utility systems is shown in Figure 2. Raw water is initially stored in specific wells and then pumped using submersible pumps to a two parallel filtration lines, each equipped with sand and anthracite filters, as well as granular activated carbon (GAC) filters. Pre-treatment steps include pre-oxidation with sodium hypochlorite (NaClO) to minimize bacterial growth and pH adjustment with carbon dioxide (CO_2) to enhance coagulant precipitation and reduce dissolved aluminum and iron, while coagulation is achieved using aluminum polychloride ($AlCl_3$). Filtration through sand and anthracite filters is performed to remove solid particles and contaminants in the water, while granular activated carbon filters further remove organic substances, chlorine, and other unwanted compounds. Both filter systems undergo periodic cleaning with water and air. The treated water is stored in two tanks, with a portion allocated for collecting cleaning water. Finally, sodium hypochlorite (NaClO) is dosed to prevent bacterial growth in the distribution network.

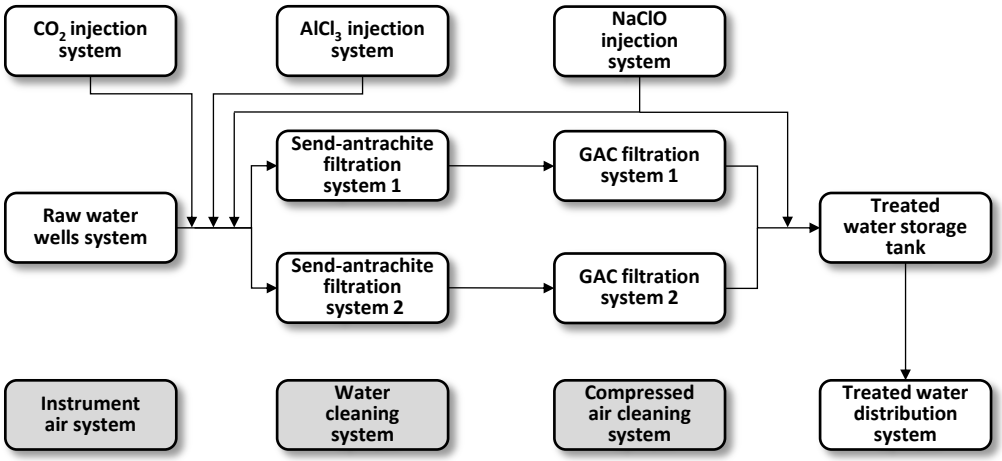

Figure 2: Block diagram of the drinking water treatment plant (grey-shaded blocks refer to utilities).

3.2 Results and discussion

The scope of analysis was set to also include operability aspects given the absence of hazardous materials processed or stored in the plant analyzed (Step 1). Each block shown in Figure 2, including utilities, was considered as a node in the analysis (Step 2). In the following, only the results obtained for node ND16, the one corresponding to the “Treated water distribution system" are presented for the sake of conciseness as POROS 2.0 application to such a number of nodes provides a large amount of output data. The main results for ND16 are reported in Table 1. A simplified P&ID of the node is reported in Figure 3.

“Cl_2 concentration out of specification” (SE01A), “Turbidity out of specification” (SE01B), “Stop of distribution pumps causing production outage” (SE02), and “Damage of distribution pumps causing production outage” (SE05) were the security events considered applicable to this ND (Step 3). They derive from a guiding list provided in the reference source of the methodology (Iaiani et al., 2023b), tailored for this specific case.

The triplets ND-SE-MA which a severity level of 2 for at least one target (cut-off criterium adopted) were selected, all reported in Table 1. What is worth noting is that the mechanisms of action to generate SE01A and SE01B require manipulations of components which belong to other nodes different from ND16 (propagation of the effect of manipulations between nodes), specifically ND03 “NaClO injection system” and ND04 “$AlCl_3$ injection system” respectively. On the contrary only manipulation of elements in ND16 are required to generate SE02 and SE05. In particular, SE05 can be obtained in two ways: by damaging the motor driver of the pumps or by exceeding operative limits. Severity vectors were estimated using the rank criteria reported in the reference source of the

methodology: while SE01A/B are deemed to generate only damages to reputation, SE02 and SE05 cause severe economic losses due to production outage combined or not with repair and/or replacement of the pumps.

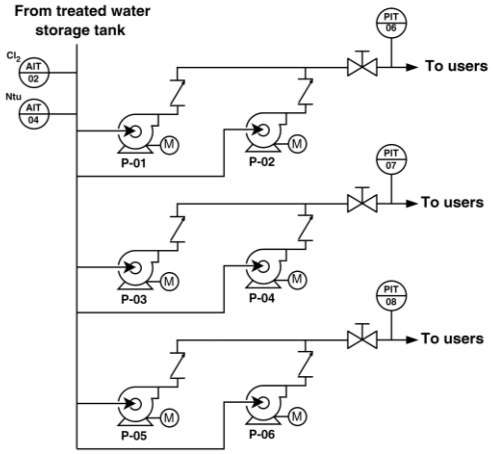

Figure 3: Simplified P&ID of node ND16 "Treated water distribution system".

Table 1: Description of the CM+APS attack actions for the ND-SE-MA triplets selected for ND16. EC: Economic loss; RE: reputation damage; PE: People damage; EV: Environmental damage; PKL: Plant Knowledge Level; CyC: Cyber Complexity.

Triplet	Description	CM + APS code	Manipulated RMCs and LCs	Effective APSs	Severity vector [EC,RE, PE,EV]	Credibility score
ND16-SE01A-MA16.1	Make Cl_2 concentration out of spec. by manipulation of ND03 "NaClO injection system"	16.1.1	Pumps PV-01/02 (increased rotational speed) in ND03	Cl_2 AIT-02	[1,3,1,1]	6 (medium PKL; low CyC)
ND16-SE01B-MA16.2	Make turbidity out of spec. by manipulation of ND04 "$AlCl_3$ injection system".	16.2.1	Pumps PV-03/04 (increased rotational speed) in ND04	Ntu AIT-04	[1,3,1,1]	6 (medium PKL; low CyC)
ND16-SE02-MA16.3	Production outage by stopping pumps in ND16	16.3.1	Pumps P-01/02/03/04/05/06 (stop of electric motor)	PIT-06/07/08	[2,1,1,1]	12 (low PKL; low CyC)
ND16-SE05-MA16.4	Production outage by damaging pumps in ND16 (driver failure)	16.4.1	Pumps P-01/02/03/04/05/06 (start/stop cycles)	PIT-06/07/08	[3,2,1,1]	12 (low PKL; low CyC)
	Production outage by damaging pumps in ND16 (exceeding operative limits)	16.4.2	Pumps P-01/02/03/04/05/06 (increased rotational speed)	PIT-06/07/08	[3,2,1,1]	12 (low PKL; low CyC)

The distribution pumps P-01/02/03/04/05/06 are the remotely manipulable components (RMC) allocated in ND16, while pumps PV-01/02 and PV-03/04 are the ones allocated in ND03 and ND04 respectively (Step 5): all these machines are driven by electric motors (the manipulative elements, ME). The fourth column of Table 1 reports the local consequences (LC) that are required in the RMCs to carry out each MA which overall form the set of combinations (CM) for each ND-SE-MA triplet (Step 6). As example, start/stop cycles of the electric motors and increased rotational speed of the pumps P-01 to P-06 are the LCs required to induce them to damage and thus stopping operations for a prolonged period of time (ND16-SE05-MA16.4), causing huge economic losses suffered by the company owning the plant, along with reputation damages due to the regional/national media

coverage of the event. The pressure indicators PIT06/07/08 present downstream the pumps in the P&ID shown in Figure 3 inform operators in the control room of possible problems in the distribution lines: therefore, they are procedural safeguards (APSs) contrasting both triplets ND16-SE02-MA16.3 and ND16-SE05-MA16.4 (Step 7). Similarly, indicators of Cl_2 concentration and of turbidity located upstream the pumps initiate responses by operators contrasting ND16-SE01A-MA16.1 and ND16-SE01B-MA16.2 respectively. After application of Step 7, the CM+APS attack actions are defined, which include all the manipulations (on RMCs and on APSs) that are required to generate the corresponding SEs. No inherent/passive safeguards (IPSs) are present in ND16 to contrast the selected scenarios: even if present, they cannot be manipulated by attackers and therefore are not part of the set of manipulations required.
Last column of Table 1 reports the credibility score obtained for each CM+APS attack action based on plant knowledge level required and cyber complexity of the set of manipulations. As it can be observed, the final scores evidence high credibility of the identified CM+APS attack actions (score ≥ 4 means credible, if ≥ 8 means highly credible according to Iaiani et al. (2023a)). This is due to the fact that a single zone is present in the network system (i.e., the BPCS) and that the plant has not specific characteristics with respect to common drinking water treatment plants (complete technical knowledge of the plant not required by the attacker). However, CM+APS attack actions corresponding to ND16-SE02-MA16.3 and ND16-SE05-MA16.4 are more credible than the others as manipulation of only components belonging to ND16 are required.

4. Conclusions

The present study focuses on investigating the effects of cyber-attacks on a drinking water treatment plant using the POROS 2.0 methodology developed by the Authors. Application to the node corresponding to the final distribution of treated water to users revealed the possibility of generating high-severe scenarios, especially in terms of production outage (by damaging vital components such as the distribution pumps) and of damage to reputation (by supplying users with water out of specification): in the latter case, implementation of automated safety systems to stop production is highly recommended. The study also investigated the credibility of the identified scenarios (function of attacker required plant knowledge and cyber-complexity of manipulations), proving that the attacks can be carried out even with scarce technical information available on the attacked system, and that not always complex patterns are required, resulting thus very credible events.
Overall, the results can be used to support cybersecurity risk assessment for drinking water treatment plants (e.g., ISA/IEC 62443 workflow) and emphasize the need for robust cybersecurity strategies to prevent and mitigate risks even in these plants where hazardous substances are not handled. Future developments involve application of POROS 2.0 methodology to food&beverage industries and the development of a quantitative model for estimating cyber risk, taking as inputs the scenarios identified with POROS 2.0.

Acknowledgments

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

References

Bajpai S. and Gupta J. P., 2018, Security risk assessment: Some techniques, Chapter In: G Reniers, N Khakzad, & P Van Gelder (Eds.), Security risk assessment in the chemical and process industry, Vol 1, De Gruyter, Berlin, Germany, 75-92.
Center for Chemical Process Safety (CCPS), 2022, Managing Cybersecurity in the Process Industries - A Risk-based Approach, Wiley, New York, USA.
Cherdantseva Y., Burnap P., Blyth A., Eden P., Jones K., Soulsby H., Stoddart K., 2016, A review of cyber security risk assessment methods for SCADA systems, Computers & Security, 56, 1–27.
Iaiani M., Tugnoli A., Casson Moreno V., Cozzani V., 2020, Analysis of Past Cybersecurity-Related Incidents in the Process Industry and the Like, Chemical Engineering Transactions, 82, 163–168.
Iaiani M., Tugnoli A., Cozzani V., 2023a, Identification of cyber-risks for the control and safety instrumented systems: a synergic framework for the process industry, Process Safety and Environmental Protection, 172, 69–82.
Iaiani M., Tugnoli A., Cozzani V., 2023b, Process hazard and operability analysis of BPCS and SIS malicious manipulations by POROS 2.0, Process Safety and Environmental Protection, 176, 226–237.
International Electrotechnical Commission (IEC), 2016, IEC 61882: Hazard and operability studies (HAZOP studies) - Application guide.