IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 15, 2021, accepted December 3, 2021, date of publication December 10, 2021,
date of current version December 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3134759

A Security-Aware Framework for Designing Industrial Engineering Processes

PANAGIOTIS DEDOUSIS[1], **GEORGE STERGIOPOULOS**[1,2], **GEORGE ARAMPATZIS**[3], **AND DIMITRIS GRITZALIS**[1]

[1]Department of Informatics, Athens University of Economics and Business, 11253 Athens, Greece
[2]Department of Information & Communication Systems Engineering, University of the Aegean, 83200 Samos, Greece
[3]School of Production Engineering & Management, Technical University of Crete, 73100 Chania, Greece

Corresponding author: Dimitris Gritzalis (dgrit@aueb.gr)

This work was supported in part by the EU-funded SPIRE-2019 FACTLOG Project (FACTLOG: Energy Aware Factory Analytics for Process Industries) under Grant GA 869951, and in part by the Hellenic Ministry of Digital Governance, through a Research Grant offered to the Research Center of Athens University of Economics and Business.

ABSTRACT Modern critical infrastructures (CI) are complex Cyber-Physical-Systems (CPS) that tightly integrate physical processes with information and communication technology components. Numerous safety mishaps and security attacks in such systems have demonstrated the need to ensure their safety and security from early design stages. Research on CPS has mostly focused on securing existing, implemented industrial systems, while safety and security consideration during the design stages of modern industrial infrastructures has largely gone unnoticed. In this paper, we present a framework that extends previous, preliminary work on the integration of security in industrial engineering design practices, and provide an algorithmic approach that effectively reduces risk during industrial system design lifecycles. We achieve this by analyzing flows of materials and information related to physical processes using three steps: (1) Identifying critical components and flows, (2) prioritizing flows based on their ties to high risk and importance in terms of dependencies, and (3) classifying system components based on their influence on the overall industrial system. To do that, we utilize (i) material flow networks (MFN) for modelling/designing the physical system (ii) dependency risk graphs for analyzing networks dependencies and assessing the system, in terms of risk, (iii) graph minimum spanning trees, and (iv) network centrality metrics. To evaluate our approach, we model and assess the production chain corresponding to an oil refinery plant's Liquefied Petroleum Gas (LPG) purification process. Preliminary findings demonstrate the complex dependencies between cybersecurity vulnerabilities and system safety.

INDEX TERMS Security by design, industrial engineering, risk analysis, risk assessment, risk mitigation, dependency risk graphs, CPS dependencies, graph centrality.

I. INTRODUCTION

CIs are cyber-physical systems consisting of physical processes, equipment and other components connected over information and communication technologies (ICT). ICT allows for system control and monitoring, thus providing functionality and process optimization across a wide range of industrial applications, improving the CI operation and the provided services [1]. ICT systems depend on physical components, while Industrial Control Systems (ICS), such as SCADA (Supervisory Control & Data Acquisition) systems, require data processing functionality to manage control and monitoring operations [2].

Continuous advances and progress in ICT led to novel CPS implementations that integrate sophisticated and complex ICT systems in CI, resulting in tighter integration between physical processes and the cyber domain. On the downside, this allows for new attack vectors to emerge, as seen in multiple security incidents such as Stuxnet, Maroochy Water Breach [3], [4]. Such security threats can impact the system's safety and reliability; therefore, security and safety must be considered as depended properties during the design phase of an ICS. However, safety and security are often treated separately, as the complexity of modern CPS makes it challenging to model and assess their interconnected nature [5], [6].

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

To overcome this inconsistent practice, engineers and security experts must focus their modelling efforts on ensuring that both security and safety requirements are considered during the design stage of CPS [7].

Authorities worldwide address the safety and security of CI as a high-priority issues [8]. The EU supports this through the NIS Directive [9], while the US has published a specific directive solely to protect CI [10]. Even though authorities have long identified the risks behind cyberattacks on CI and despite the numerous advances in CPS protection, still, to our knowledge, there is little work on how to consider both safety and security principles in industrial critical infrastructure engineering design. Investing in system architecture and integrating security and safety requirements early in the design stage is far more efficient and cost-effective than funding the protection of vulnerable and insecure systems [11], [12].

A. CONTRIBUTION

We propose a system design framework for CPS that integrates both security and safety. Critical industrial infrastructures are production systems with complex production chains. In our work, these production chains are represented as networks of productive activities characterized by flows of resources, i.e., physical flows of materials and energy and flows of monitoring and control information. Physical flows are subject to availability requirements and constraints of the output capacity of the production system. Similarly, monitoring and information availability and integrity are required to ensure the system's output. To that end, we employ risk analysis and dependency analysis to assess a critical industrial infrastructure production chain.

This paper extends the work presented in [13] concerning security integration in industrial engineering design practices. The presented method builds on a previous work that models CI production chains as material flow networks. Material flow networks (MFN) are directed graphs consisting of vertices that represent the location of material and energy transformations (processes) or storage of a production chain, and edges signify material and energy flows between them [14], [15] (see the discussion in Section IV.A for more on this). Engineers primarily utilize MFN to analyze system flows for optimization purposes based on multiple criteria (e.g., cost, environmental and social impact) [16]–[19]. Authors in [13] model and analyse material flow networks to detect and identify high-risk channels (paths) and critical components (flow network nodes) based on their overall effect on the system. The presented approach utilizes this technique to (i) model processes and flows in production chains of CI into MFN (ii) build an efficient flow model able to distinguish and map only the required flows and processes (essential in terms of risk) of an MFN into a dependency graph, simplifying calculations and improving applicability, (iii) calculate the likelihood value for each flow network node of the mapped dependency graph, considering both its failure rate and the lack of a required resource. In addition, it extends the previous work by introducing minimum spanning trees (MST) and centrality metrics to efficiently identify and prioritize high-risk flows and flow network nodes for risk mitigation. More importantly, it presents a clear roadmap/workflow to guide the risk mitigation efforts of engineers and security experts during the design stage acknowledging specific risk goals and system requirements.

To evaluate our approach, we assess a part of the production chain corresponding to the Liquefied Petroleum Gas (LPG) purification process of the TUPRAS oil refinery plant [20]. The material flow network data for the TUPRAS use case was provided by the EU-funded SPIRE-2019 FACT-LOG project [21]. In summary, our paper contributes the following:

1. An improved modelling approach that maps and converts the assets and the interdependencies of a material flow network into a risk dependency graph based on the existing production chain topology of a CI.
2. An improved risk calculation methodology that depicts a threat's probability of disrupting a CI asset based on a noisy-OR mode.
3. High-risk channel (path) identification and prioritization utilizing dependency risk analysis
4. Critical flow identification and prioritization utilizing a minimum spanning tree (MST) algorithm.
5. Critical flow network node identification and prioritization utilizing network centrality metrics.
6. A framework that provides a clear roadmap to guide and assist engineers and security experts' risk reduction efforts during the design stage of CPS.

B. STRUCTURE

The rest of the paper is organized as follows: Section II discusses related work and compares CPS protection methods. Section III describes the proposed framework. Section IV describes the fundamental building blocks of the framework. Section V discusses the methodology implementation in a real-world example and presents our findings to validate the methodology. Finally, the conclusion discusses paper results and potential future research in section VI.

II. RELATED WORK

Various methodologies are used to analyse CPS of CI security or safety and evaluate the different dimensions involved in the factors that affect CI operation and provided services [22]. The main goal of such high-level methodologies is to analyse and evaluate threats and multi-dimensional impacts of disruptive incidents involving CI in multiple sectors [23], [24].

Traditional risk assessment methodologies usually focus on vulnerabilities on IT systems of CI [25]–[27]. These assessments are performed on already established and functioning systems and primarily result in added layers of cybersecurity on top of existing systems. However, the ICT sector is not aided with security standards to the same extent as the IT domain. The ISA/IEC 62443 family of standards targets ICS and Industrial Automation and Control Systems (IACS) security [28]. The concept to perform a security risk assessment and management is similar to what is outlined in the ISO 27000 family [29]. To that end, traditional security technologies of IT systems are adapted to protect CPS security; however, threats from the CPS cyberspace are mainly unpredictable and untenable; thus, traditional reliability theory and fault-tolerant technology cannot completely prevent the system from failures [22].

Other approaches that focus on CI primarily delve into dynamically assessing industry IT and ICT networks by evaluating the cascading failures over time between assets involved in and among different business processes [30]. Most are utilizing graphical models over the system architecture and perform risk analyses to understand ICS (i.e., PLC, RTU, SCADA) weaknesses in the industry [31], [32]. Others perform targeted, technical attacks on individual ICS systems, e.g., binary manipulation of ladder logic in PLCs, attacking actuator software, etc[33], [34]. Authors in Roy *et al.* [35] utilize attack and counterattack trees to perform qualitative and probabilistic analysis of the security status of CPS. Other approaches utilize the concept of security-by-design to provide more flexible and effective ways to secure ICT/OS solutions during software development [36], [37]. However, while addressing the risk of attacks to a CI, these approaches focus mainly on cybersecurity, ignoring the various threats and vulnerabilities of the various physical processes, components, and machines involving in the production chain of a CI.

Several simulation-based approaches have been developed to analyse and assess threats to improve security, thus the reliability and resilience of CI under attack scenarios [38], [39]. The main problem with statistical model checking is that the probability estimation becomes unfeasible for rare events. Authors in Ferrario *et al.* [40] proposed a method that utilizes a Monte Carlo simulation and a Hierarchical Graph to model ICS dependencies and evaluate CI robustness. Their modelling approach presents several similarities with our methodology, but they focus on high-level dependencies and supply chain conflicts, overlooking dependencies between physical processes in complex production chains. For instance, in modern production systems, various waste and energy recovery systems are applied that create circular dependencies in the production chain [41], [42].

Many methodologies for safety risk assessment have been developed for CPS. Risk and hazard analysis techniques can be categorized into two groups: (i) Failure-based hazard analysis and (ii) systems-based hazard analysis techniques. Failure-based techniques include Fault Tree Analysis (FTA) [43], [44] and Failure Modes and Effect Analysis (FMEA) [45], [46]. Failure-based methods focus on the identification of the effects and probabilities of single component failures omitting failures rooted in the interaction of components. System-based analysis techniques include the Hazard and Operability Analysis (HAZOP) [47], [48] and the system theoretic process analysis (STPA) [49], [50]. System-based approaches do not focus on cybersecurity; thus, it dificult to quantify risk in modern systems that rely increasingly on IT and ICT to control physical processes. Also, both software vulnerabilities and the effects of non-technical influences on the system over time are very hard to measure at design time [51].

Several approaches attempt to safeguard both safety and security in ICS as addressing both safety and cybersecurity is paramount to modern CPS' smooth and trustworthy operation [52]. Others adopt the concept of system-of-systems to address security, reliability, and robustness in CI [40], [53], [54]. However, although system-of-systems analysis provides a comprehensive top-down overview of the environment in which a CPS operates and how risk propagates to and from the system, it is fraught with uncertainty about how constituent CI systems operate and function. Furthermore, the challenge with most of these methods is that they focus only on the physical system from a safety perspective and not on the complete communication system [55].

System optimization is commonly used during the design stage of a production system [56]. To that end, traditional risk and safety assessments are used to optimize and create a secure and safe system at the early stages of the system lifecycle [57]. The main issue with these approaches is that they focus explicitly on cybersecurity or safety threats, neglecting the relationship between security and safety. From an engineering perspective, the concept of optimization in system design is not new, as material flow analysis (MFA) and material flow networks (MFN) are utilized during the design stage to optimize the model system based on multiple criteria (e.g., cost, environmental and social impact) [16]–[19]. However, most of them focus on cost-effectiveness and do not consider the security or safety perspective of CPS in critical industrial infrastructures [58], [59]. Other techniques implement security-by-design during the implementation stage by selecting certified components based on specific cybersecurity standards [60]. Others implement safety-by-design by selecting components considering various safety factors [61]. To that end, security and safety by design should go beyond selecting the individual system components and how they are secure or safe based on their design.

Similar to the work in [13], the proposed approach focuses on individual critical industrial infrastructures (like energy corridors for oil and gas supply, water, and waste treatment plants). In [13], authors embed security-by-design concepts in industrial critical infrastructure engineering by integrating a security risk assessment process into engineering design practices. In our approach, we build on this previous work by introducing an algorithm for detecting high-risk flow network nodes and flows to this risk assessment process. These high-risk nodes and flows are then used as input to a risk mitigation algorithm that extends previous risk assessment. We incorporate this entire process in a novel framework that provides an algorithmic roadmap for introducing risk mitigation decisions to industrial engineers during the design phase of CPS. Similar to the work in [13], we utilize MFN [17]–[19] to model the underlying system. We employ

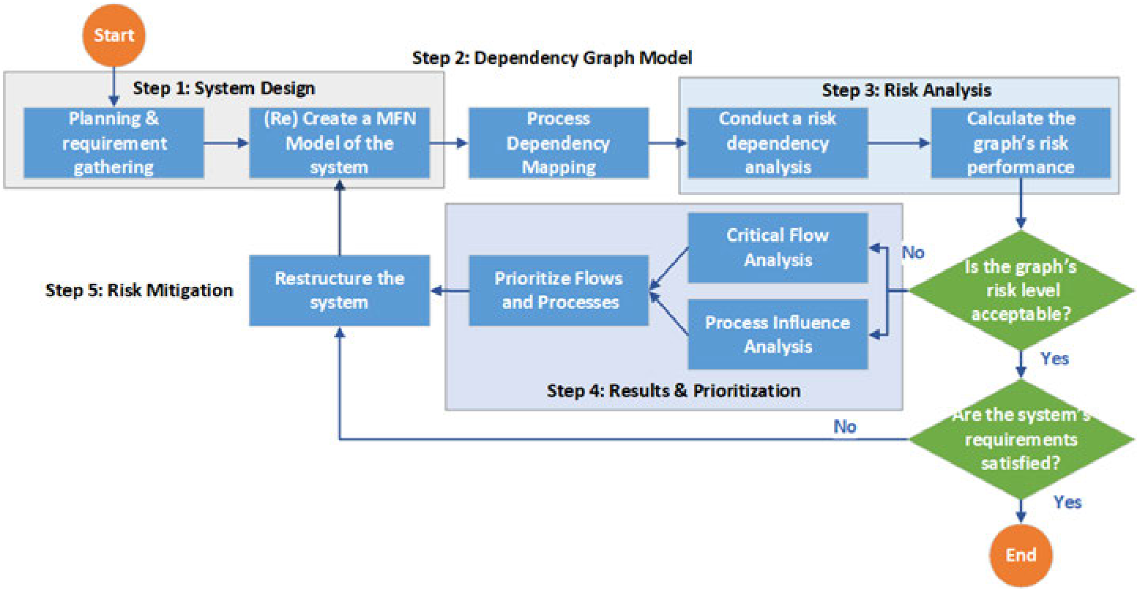

Start
Step 1: System Design
Planning & requirement gathering
(Re) Create a MFN Model of the system
Step 2: Dependency Graph Model
Process Dependency Mapping
Step 3: Risk Analysis
Conduct a risk dependency analysis
Calculate the graph's risk performance
Is the graph's risk level acceptable?
No
Yes
Critical Flow Analysis
Process Influence Analysis
Prioritize Flows and Processes
Step 4: Results & Prioritization
Restructure the system
Step 5: Risk Mitigation
Are the system's requirements satisfied?
No
Yes
End

FIGURE 1. High-level overview of the proposed framework.

traditional risk assessment methodologies to evaluate the ICT processes [25]–[27] and hazard analysis techniques such as FMEA, FTA, HAZOP, STPA to assess the physical processes of the modelled system [43]–[50]. At the same time, we utilize a similar methodology with [30], [62], [63] to model dependencies between the different components of the modelled CPS and assess the cascade risk due to possible disruptions considering an all-hazardous approach.

Previous risk analysis approaches focus mainly on dependencies and study cascading failures between individual CI or between discrete components of corporate IT networks.

However, these approaches do not yet address cascading failures between cyber and physical components inside individual infrastructures, nor they are able to analyse their subliminal attack paths. Such attacks can cause internal cascading malfunctions to equipment, which may lead to a chain reaction affecting other components due to erroneous data reports, injection or corruption of sensor information and actuator orders in communication channels, or unavailability of service. We tackle this gap by presenting a risk management framework that can be embedded inside industrial design processes. To achieve this, we analyse the interconnected components of a CI considering conditional probabilities of component failure and supply disruption due to threat manifestation in order to calculate the cascade risk of attack paths.

Our modeling approach resembles solutions that use the top-down method to analyze complex interdependencies in individual CI, similar to the work in [40], [64]. In [40], the authors consider engineered, physically networked (energy, transportation, information, and telecommunication) CI and their interconnected components to quantitative evaluate CI robustness. However, their approach neither provides metrics for cascade risk evaluation nor suggests mitigation controls to improve the CI robustness. In our approach, we suggest risk mitigation controls and guide expert efforts to reduce the risk of the individual CI by utilizing MST and centrality metrics following the techniques proposed in [65] for automated IT network risk reduction.

III. FRAMEWORK

The proposed framework follows the standard development system lifecycle incorporating security and safety criteria into the design phase. It spans all system processes, as security risks will need to be identified as early as the design phase and addressed accordingly.

The framework aims to assist designers in understanding the potential impact of compromised components and identifying and prioritizing weak points for risk mitigation. Fig. 1 summarises the steps and workflow of the proposed roadmap.

1. First, following the standard development lifecycle, engineers should discover system requirements and create a system model for evaluation during the design phase.
2. Next, system engineers with security experts should evaluate the modelled system (in terms of risk);
   a. if the risk level is acceptable and the system requirements are satisfied, the development lifecycle can continue. However,
   b. if the risk levels are above a threshold value, the system must be analysed to prioritize flows and processes for risk mitigation.

The threshold parameter is subjective, as per real use-cases when setting up industrial processes; a decision-maker can define the parameter based on the critical industrial infrastructure-under-design specific characteristics. Risk mitigation measures, in our case, include the addition or subtraction of network nodes and flows or the replacement of a node with similar functionality but a lower failure rate and impact value.

Each step of our methodology utilizes a set of mapping procedures and algorithms, where each one provides some insight on the critical industrial infrastructure production chain under analysis and outputs information to be used as input by the following step. Below we present the fundamental steps of our framework:

I. **Process dependency mapping**: We identify, input, and map an industry's cyber and physical processes and process flows (MFN) into a dependency graph

II. **Process dependency risk analysis**: We assign failure probabilities and impact values for each node. The algorithm pre-computes all n-order dependencies using the process dependency graph. Then for each dependency chain, outputs the cumulative dependency risk of each attack path. Finally, the algorithm calculates the overall risk for the mapped MFN.

III. **Critical flow analysis**: The tool produces alternative graphs with minimum risk (MST), maintaining process connectivity, using the process dependency graph, and computes the removal rates for the removed high-risk dependencies, thus identifying and prioritize critical flows for risk mitigation.

IV. **Process influence analysis**: The tool pre-computes the centrality metric values for each node using the process dependency graph and highlights the maximum values, thus identifying critical processes for risk mitigation.

V. **Steps beyond the scope of this work**: The development of mitigating measures and the improvement of the system is not covered in this paper. However, it is intended that the results of this work are used as input for those steps. Also, in this work, we do not address the initial design of the system (in terms of defining the architecture, the processes, system parameters, etc.) and its requirements both in terms of quality and quantity. However, strictly defined system requirements are crucial to ensure that the selected mitigation measures do not affect the industry's qualitative and quantitative objectives.

In the following section, we discuss in detail the building blocks that compose our framework.

IV. BUILDING BLOCKS

This engineering design methodology uses six (6) building blocks:

1. A Material Flow Network (MFN) method for modelling material, energy, and informational flows in production chains of critical industrial infrastructures as a Material Flow Networks based on MFN principles.
2. A modelling method that maps and converts the flow network nodes and flows of a material flow network into a risk dependency graph based on the existing production chain topology of a CI.
3. A risk calculation methodology to estimate the likelihood of a threat disrupting the system components' operation.
4. A multi-risk dependency analysis methodology for assessing the risk of the graph's dependency paths and the graph's overall risk.
5. A minimum spanning tree (MST) algorithm for critical flow identification and prioritization.
6. Network centrality metrics to identify influential critical flow network nodes.

Each building block is briefly presented below.

A. MATERIAL FLOW NETWORK MODELLING

Material Flow Analysis (MFA) is a systematic and analytical method for the mapping of flows and stocks of materials within a system defined in space and time. It aims to connect the sources, the pathways, and the sinks (targets) of materials in the system [66]. In MFA, Material Flow Networks (MFNs) serve as reference models (templates) for developing more refined/optimized models in the same domain. MFN are based on the popular Petri Net methodology for specifying concurrent systems in Computer Science and have been transferred into the Environmental Sciences [15]. The use of MFN allows the representation of material flow systems as directed graphs where vertices represent single manufacturing steps or places where materials and energies are processed/transformed or stored. Graph vertices are linked by edges that correspond to the material flows within the system [14], [15]. Engineers primarily utilize MFA and MFN for process optimization and eco-balance (the process of efficient utilization of material resources and energies and balance of environmental impacts) [16]–[18].

In our approach, we follow the principles of MFA and utilize MFN to model processes and material and energy flows in production chains, similar with the work in [16]. We model flow networks as graphs with four types of nodes: (i) processes, (ii) junctions, (iii) input, and (iv) output nodes. These are connected via links (Fig. 2). Single activities in which resources (material, energy, and information) are processed and converted/transformed into other resources are referred to as *Processes*. *Input Nodes* are the initial sources of resources flowing towards processes and represent different external resource suppliers (e.g., industries, CI). *Output Nodes* are the final recipients of resources flowing from processes, and they represent various external resource receivers (e.g., the environment) or consumers (e.g., industries, households, CI). Finally, *Junctions* serve as storage nodes for network resources, connecting processes and serving as output nodes and input nodes for other processes. For all intents and purposes, processes and junctions represent the assets of a modelled infrastructure. Resources can flow between nodes via *Links*, constituting a mode of transport (e.g., pipes, cables, roads, ships). Such flows between nodes describe the rate at which resources are consumed (input flows) and produced (output flows).

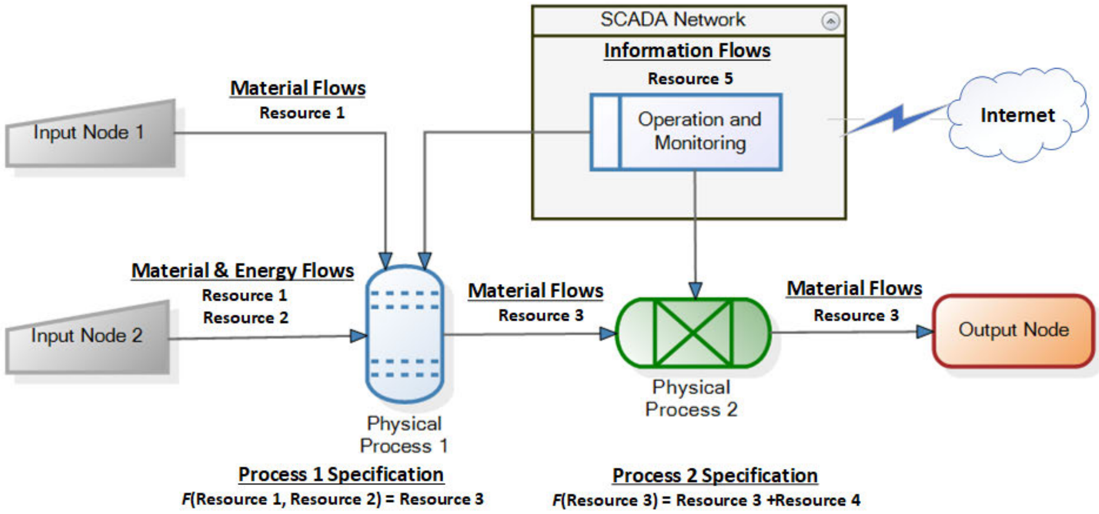


FIGURE 2. Graphical representation of a MFN capturinh the correlation between the cyber and physical components of a production system.

For modelling purposes, we characterized input flows as regular or backup, determining the consumption lifecycle (i.e., a regular flow transfers resources continuously from parent to target node). Also, input or output flows are assigned to the same link if they share the same transport modes. Fig. 2 showcases a demo MFN demonstrating the various types of nodes and the flow of resources between them.

1) MODELLING PROCESS SPECIFICATION

A critical step in the modelling process is process specification. This includes the definition of input and output resources and the relations between input and output flows and process parameters (i.e., environment temperature). Physical process nodes are the principal entities and comprise the majority of flow networks. They depict activities that begin with a set of resources (input) and end with new or processed resources (output). Thus, modelling system physical processes help researchers investigate and analyse various dangers posed by human actors (outsiders, insiders), natural disasters, and hazardous situations.

In our approach, as an exception to the preceding concept, processes can also describe industrial automation control and monitoring activities of the system that supply and receive monitoring and control data from and to connected processes or junctions. From a hierarchical perspective, we should note that automation control and monitoring activities are essentially subprocesses, at a different level from the physical processes that operate on top of them, providing crucial functionality. Modelling the automation control and monitoring activities as a part of the overall system allows us to study and analyse various types of cyber threats that leverage vulnerabilities of monitor and control information systems and the impact they pose to the physical components of the modelled system.

Using the methods provided, we can create a model that represents an actual CI production system. If all processes are accurately specified, and more importantly, their input and output flows, then and only then we can have a holistic view of the system and understanding of the various and complex interactions/dependencies between the physical and cyber elements of the CI production chain so that we can evaluate the risk level of the CI.

B. MODELLING DEPENDENCY GRAPHS

To analyse and assess the flow network nodes' risk and evaluate the infrastructure's overall risk, we must first pre-process the flow network and map MFN nodes and flows into a risk dependency graph.

In the pre-processing stage, we mark modelled junctions used to collect multiple flows into one as ignorable because they are used solely for mass-balance calculation purposes and do not reflect physical objects in the real world. Also, for each process, we select the required resources (Primary flows) for its operation. Engineers based on the system specification requirements may model secondary flows for simulation purposes. For instance, a water treatment process requires water and chemicals for its purification, so an engineer will model both water and possible impurities it carries as input flows to the process.

After removing unnecessary junctions and secondary flows, we map all material flow network nodes (input, output, junction, and process nodes) as potential failure nodes, together with all flows between them, into a risk dependency graph.

Dependencies are modelled in directed, weighted graphs $G = (V, E)$ where the nodes V represent the possible failure nodes of the system and edges E represent the dependencies between them. The weight of each node (i.e., process, junction, input/output node) quantifies the estimated dependency risk of flow network node B on resources provided by flow network node A. This weight derives from the dependency between the flow network nodes. Weight calculation is presented below in Section IV.C. The resulting graph depicts the movement of resources from one possible failure node to another as input and output dependencies.

C. RISK ANALYSIS

In this section, we look at how to measure risk, which factors contribute to risk in the cyber-physical model (including those unique to critical industrial infrastructures). Risk factors and likelihood evaluation are thoroughly discussed in the following subsections.

1) RISK MODEL FACTORS

A risk is the degree of possible failure that may occur in an established process, and risk assessment is one of the critical activities in the risk management process. The standard reference of risk as a cybersecurity assessment metric is the following Risk $=$ Likelihood $*$ Impact. Assessing risk means identifying the threats and then determining the likelihood and impact [25], [26], [60], [67]. We focus on external/internal threats in input resource supply, along with process availability risks. Such threats arise from malicious, natural, or accidental events. These high-impact events are unexpected, can cause a severe dysfunction of an internal process or the supply of a resource, and, more importantly, they can propagate down the production chain. Similarly, with Adenso-Diaz *et al.* [68], we do not differentiate between disruption types but rather consider disruptions in general and their effect on the production line. Thus, each node in the flow network is potential failure node that is either entirely disrupted or fully operational. This binary approach is a typical way to model disruption of resources supply [69], while it can be used to simulate disruptions in the field of CIP [70].

Due to the cyber and physical aspects and threats of the CPS, as indicated in Fig. 3, we utilize traditional risk and hazard analysis methods, such as ISO/IEC 27001, HAZOP, FMEA, to estimate the impact of cyber and physical threats of each flow network node in the cyber-physical system. Impact as a metric depicts the magnitude of harm due to the loss of availability or integrity of a flow network node (i.e., process). For example, the loss of a CI process due to a threat realized affects all dependent CI processes in the production chain, thus the system's availability and integrity. In many cases, a compromised CI process could result in significant loss of life, casualties, material harm, environmental damage, and public service disruption.

2) LIKELIHOOD CALCULATION

In this stage, we calculate and assign a likelihood value to each node of the mapped risk dependency graph based on the initial flow network model flows. This value indicates how probable it is for a threat to disrupt the operation of a flow network node (i.e., a process, junction, or input/output node) by impeding its activity or disrupting the supply of one or more required resources. An individual flow network node N (a piece of equipment) can have two states: either failed (N) or functional ($\bar{N}$); similarly, an input resource R is either unavailable (R) or available ($\bar{R}$). We should note here that the operation of flow network depends on its own ability to function and the availability of its required resources. As such, the probability of the event (x) the operation of a single node N with required resources R to be disrupted given the node N and the required resources R are in the state u is $P\,(x \mid u)$. The probabilities $P\,(x \mid u)$ are called risk parameters, and with binary states, there are 2^{n+1} parameters to be defined for a flow network node with n input resources.

For a network flow graph G with V nodes and E edges/transfer links, the likelihood of a disruption L_i for the operations of a node $i \in V$ is calculated using (1).

$$L_i = \sum_{\forall u} P\,(x \mid u) \tag{1}$$

where $P\,(x \mid u)$ denotes the conditional probability for the event (x) the operation of node i to be disrupted given the node i and its required input resources are in state u.

To evaluate the relationship of these disruptions between network flow nodes and input resources, we utilize a noisy-OR model [71] similar to [72], [73]. The Noisy-OR model assumes the independence of causal influences among a flow network node and its required input resources [71]. This assumption provides a logarithmic reduction in the number of parameters required; for a flow network node with n input resources, there are $n + 1$ independent parameters. By minimizing the number of network parameters, we improve the implementation process for real-world applications. This way, we reduce the computational and modelling challenges that MFN models introduce.

To this end, the state of node i (i.e., failed N or functional $\bar{N}$) depends on its failure probability a_i. The failure probability $a_i, \forall i \in V$ estimates how likely it is for a node i (a piece of equipment) to fail individually at some point in the future, as such the probability for node i to be operational is $(1-a_i)$. Similarly, the availability state of a required resource j (i.e., available $\bar{R}$ or unavailable R) depends on its disruption probability FLR_j. The required input resource disruption probability FLR_j depicts the probability of a required input resource j of node i to be unavailable, as such the probability for the required resource j to be available is $(1-FLR_j)$. Table 1 demonstrates the likelihood L_ι calculation for a node i with one required resource j based on the different states u. We should note here that the likelihood L_ι is calculated recursively as a flow network node may have more than one required resource. Also, for nodes without any input resources, we have $L_i = a_i$.

The availability of an input resource depends on its regular and backup flows from supplier flow network nodes. To that end, the disruption probability FLR_j of a required resource j is calculated based on the disruption probability LR_j due to

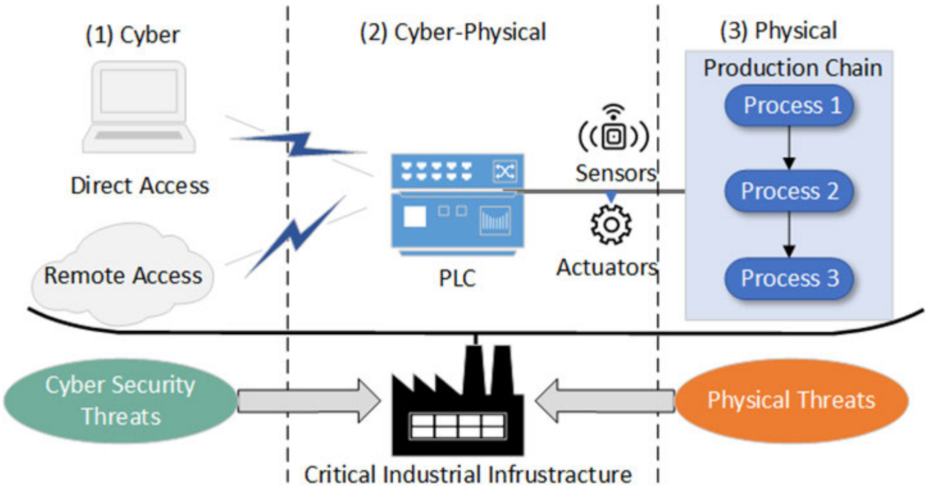


FIGURE 3. Cyber and physical aspects and threats of a cyber-physical system in a critical industrial infrastructure considering a typical PLC scenario. The cyber aspects (1) are the cyber interactions with the PLC subject to cyber threats (control centers). Cyber-physical aspects (2) are those that connect cyber and physical aspects. (i.e., PLC, actuators, and sensors). Finally, the physical aspects (3) are the physical objects subject to physical threats and hazards that need monitoring and control (i.e., the physical processes, machines, storage tanks).

TABLE 1. Node with one resource flow likelihood calculation example.

States u	$P(\mathrm{x}\|u)$
$u_1 = \{\bar{N}, \bar{R}\}$	$(1 - a_i)(1 - FLR_j)$
$u_2 = \{\bar{N}, \mathrm{R}\}$	$(1 - a_i)FLR_j$
$u_3 = \{N, \bar{R}\}$	$a_i(1 - FLR_j)$
$u_4 = \{N, \mathrm{R}\}$	$a_i FLR_j$

regular flows, and the number of backup flows B_j utilizing (2). We should note here that we assume that the backup flows will always be available on demand.

$$FLR_j = \frac{LR_j}{B_j + 1} \tag{2}$$

The likelihood LR_j reflects the availability probability of an input resource j from supplier flow network nodes that share the resource demand through regular flows. To that end, disruption probability LR_j is calculated based on the failure probability a_i of the resource supplier flow network nodes utilizing (3). If a required input resource j is supplied by only one flow network node i then $LR_j = a_i$

$$LR_j = \sum_{\forall v} P\left(y \,|\, v\right) \tag{3}$$

where $P\left(y \,|\, v\right)$denotes the conditional probability for the event (y) the availability of a required input resource j to be disrupted given the resource j supplier flow network nodes are in state v.

To this end, the state of a supplier node i (i.e., failed N or functional $\bar{N}$) depends on its failure probability a_i, so the probability for a supplier node i to be operational is $(1-a_i)$. The calculation of the disruption probability LR_j for a required input resource j of a process node P that it is supplied from two flow network nodes $\{N1, N2\}$ with regular flows is demonstrated in Table 2. We should note that the disruption probability $LR_j$is calculated recursively as a required input resource may have more than two supplier nodes.

TABLE 2. Resource with regular flows likelihood calculation example. Note that due to the noisy-or assumption, there is no need to specify separately the probability for resource j being unavailable (event y) if both $N1$,$N2$ are disrupted (state u_4).

States u	$P(\mathrm{y}\|\mathrm{v})$
$u_1 = \{\overline{N1}, \overline{N2}\}$	$(1 - a_{N1})(1 - a_{N2})$
$u_2 = \{\overline{N1}, N2\}$	$(1 - a_{N1})a_{\mathrm{N}2}$
$u_3 = \{N1, \overline{N2}\}$	$a_{N1}(1 - a_{N2})$
$u_4 = \{N1, N2\}$	$a_{N1}a_{N2}$

The disruption (failure) probability a_i for a flow network node i (system component) can be estimated based on the reliability of the component. The reliability of a system component (machine) is its probability of performing its function within a defined period with certain restrictions under certain conditions. The following databases are a subset of the best-known data sources for reliability and failures rate information for machines and processes: OREDA, GIDEP, TUD database, SRDF, European Industry Reliability Databank, CORDS, CCPS [74].

An alternative data source is failure analysis methods such as FTA or FMEA performed on similar systems. These methods are commonly used in safety and reliability analysis to understand how systems can fail and determine (or get an estimation of) event rates of a safety accident or a particular system level (functional) failure.

D. DEPENDENCY RISK ANALYSIS

After having calculated the likelihood of disruption for each node, the methodology moves on to assess the risk of first-order dependencies. For each a graph node A with risk R_A

all output edges to receiver nodes have a cascade risk R_A. Potential disruption to a component of a CI is transferred from the previous connection to the next, where the disturbance of a required input resource, regardless of the cause, may propagate to the dependent components in the production chain. To calculate and assess the nth-order cascading risks propagated in a series of components, we use the following method that utilizes a recursive algorithm based on [62], [63]. Given $A_1 \rightarrow A_2 \rightarrow \ldots \rightarrow A_n$ is an nth-order dependency between n networked components, with weights $R_{i,i+1} = L_{i,i+1} I_{i,i+1}$ corresponding to each first-order dependency of the path, then *the cascading risk exhibited by A_n for this component dependency path* is computed utilizing (4).

$$R_{1,\ldots,n} = L_{1,\ldots,n} I_{n-1,n} = (\prod_{i=1}^{n-1} L_{i,i+1}) I_{n-1,n} \tag{4}$$

where L_i is the disruption probability of a flow network node, as calculated by using (1).

The cumulative dependency risk is the overall risk exhibited by all the components in the sub-chains of the nth-order dependency. If $A_1 \rightarrow A_2 \rightarrow \ldots \rightarrow A_n$ is a chain of asset dependencies of length n then the *cumulative dependency risk*, denoted as $CR_{1,\ldots,n}$, is defined as the overall risk produced by an nth-order dependency:

$$CR_{1,\ldots,n} = \sum_{i=1}^{n} R_{1,\ldots,i} = \sum_{i=1}^{n} (\prod_{j=1}^{i-1} L_{j,j+1}) I_{i-1,i} \tag{5}$$

Finally, using the total number n of all asset sub-chains (possible asset dependency paths) and their cumulative dependency risks, the methodology calculates the graph's overall dependency risk G_r as the *sum of the cumulative dependency risk for each nth-order dependency* in the graph:

$$G_r = \sum_{i=2}^{n} CR_{1,\ldots,n} \tag{6}$$

E. MINIMUM SPANNING TREE ALGORITHM

The algorithm then utilizes minimum spanning trees (MSTs). MSTs are commonly used to find approximate solutions for complex network problems such as the Traveling Salesman [75], [76]. A spanning tree T of a weighted undirected graph G is a connected subgraph of G such that (i) T contains every node V of graph G, and (ii) T does not contain any cycle. A cycle is a graph path in which the first node corresponds to the last. A minimum spanning tree (MST) has the minimum total weight [77]. An MST of a weighted undirected graph can be found by greedy algorithms such as those described in [76], [78].

We create all possible MSTs' of the dependency graph. This way, we can compute and output the removal rates of each dependency (Removed Dependencies Report*). If an edge is continuously removed from many MST, this has significant ties to high risk and importance* in terms of dependencies and must be considered for mitigation measures/actions. To that end, we utilize MSTs to *detect and prioritize high-risk material and informational flows in MFNs for mitigation measures*.

The MST graph's calculations are performed on the dependency graph using an implementation of Prim's algorithm [78]. Starting from a critical asset node (e.g. servers), any adjacent node with the smallest weight edge is selected and added to the tree. This process is repeated, always choosing the minimal-weight edge that joins any connected node not already in the tree. When there are no more nodes to add, the tree is a minimum spanning tree. Obviously, the resulting graph $G' = (V, E')$ is undirected and non-circular, since it is based on Prim's algorithm, where $E^{'} \subseteq E$.

If graph cycles exist and have multiple dependencies of equal weight, many alternative MSTs can be produced. Note that each MST of the graph contains the same number of dependencies, and this number is guaranteed to be the smallest possible that retains the graph's connectivity. However, since the produced MST is an undirected graph, there is no guarantee that it also contains the minimum number of directed paths that can retain the flow network nodes connectivity once the directions of the necessary edges are applied. Also, the removed dependencies from the MST production represent flows of material and information required for nominal system operation. As a result, MSTs are not applicable redesigns and cannot be utilized as MFN restructures themselves for risk mitigation purposes.

F. CENTRALITIES METRICS

At this stage of our analysis, the presented method also utilizes centrality metrics on the produced risk dependency graph. Centrality metrics are widely used in network analysis and flow management [79]–[81]. In graph theory and network analysis, centrality metrics attempt to quantify the position of a node in relation to other nodes and to estimate the relative importance of a node within a graph.

In a risk dependency graph, centrality metrics can be used as additional criteria to identify the flow network nodes that significantly affect the critical risk paths of the graph. Such nodes are suitable to consider when prioritizing mitigation controls. Therefore, if appropriate mitigation controls are applied to these nodes, multiple cumulative dependency risk chains can be reduced, thus lowering the overall graphs risk [82], [83].

Note that different centrality metrics capture different aspects of network topology and thus describe different types of node influence. For example, previous comparative research on centrality metrics on dependency graphs allowed us to opt for two different centrality metrics to identify the most influential nodes [82].

1) BONACICH CENTRALITY

The Bonacich (eigenvector) metric [84] measures the centrality of a node in a network. It is calculated using the following equation:

$$c_i(\alpha, \beta) = \sum\nolimits_j (\alpha - \beta c_i) R_{i,j} \tag{7}$$

where α is a scaling factor, β reflects the extent to which centrality is weighted, R is the node adjacency matrix, j is the identity matrix and i is the matrix of ones. An adjacency matrix is a $N \times N$ matrix with each element assigned a value of 1 if an edge exists between the corresponding nodes and 0 otherwise.

A flow network node with high Bonacich centrality *is adjacent to flow network nodes with very high (or very low) influence*, depending on whether the parameter β is greater or less than 0 [84]. In a risk dependency graph, nodes with high eigenvector centrality (when $\leq$ 0) are of particular interest because they are connected to other important nodes with high connectivity. Their influence is proportional to the total risk of the first-order dependencies that affect it.

2) CLOSENESS CENTRALITY

This metric quantifies the central or peripheral placement of a node (asset) in a two-dimensional region based on geodesic distances. It is defined as:

$$C_c(x) = \sum\nolimits_{\forall x \in V(G)} d(x, y) \tag{8}$$

where $d(x, y)$ is the average shortest path between node x and any other node in the graph [85]. Closeness centrality captures *the average distance between a node and every other node in the graph* and assumes that nodes can only pass influence to their existing edges. The normalized form of the closeness centrality represents the average length of the shortest paths instead of their sum and can be defined as:

$$C(x) = \frac{n-1}{\sum_{i=1}^{n} d(x_i, y_j)} \tag{9}$$

where $d(x_i, y_j)$ is the distance between nodes x and y [86].

A flow network node with high closeness centrality has short average distances from *most other flow network nodes in a graph*. Also, if a flow network node has high closeness centrality, it is in a position to propagate disturbance quickly.

In a dependency risk graph, nodes with high closeness centrality tend to be part of many dependency chains. In most cases, these nodes initiate fast cascading effects throughout a network [82] since cascading effects tend to affect relatively short chains [87]. The closer a node is to the initiator of a cascading event, the greater its effect is on the cumulative dependency risk because the likelihood of its outgoing dependency would affect all the partial risk values of subsequent dependencies (edges).

G. ANALYSIS OUTPUT INFORMATION

The proposed methodology overall output comprises from:

- metrics that assess the performance of the flow network (i.e., the flow network graph overall dependency risk, the top and average cumulative dependency risk, the number of attack paths),
- an identification of the most critical (in terms of risk) dependencies (flows) and paths between flow network nodes, and
- an identification of the most critical flow network nodes based on influence and presence in dependencies and paths.

The performance metrics and the identified high-risk flow network nodes and flows are then used as input to the risk mitigation algorithm as described in Section III. The risk mitigation algorithm is calculated iteratively, producing alternate redesigns of the original flow network until the risk level of the system is deemed acceptable. The outputs mentioned above are recalculated on each iteration, helping the experts decide their mitigation actions in order to redesign the system. The final output is an optimized system design model based on the parameters and the decisions made by the experts during the analysis.

V. EVALUATION

A. TOOL IMPLEMENTATION

The methodology was developed as a distributed application (entitled "Process Simulation Modelling tool – PSM), including a desktop application and a web application. The desktop main application front-end and back-end are developed and implemented in the.NET framework using C#. The main application handles the modelling functionalities and the preliminary risk analysis. The web application back-end is developed in Java Spring using the Neo4j graph [88] and handles the risk dependency analysis. The desktop application front-end communicates and interacts with the web application back-end through an application programming interface (API). All the experiments were performed using a computer with an Intel Core I7, 4.6 GHz processor with eight cores, and 16 GB RAM.

B. CRITICAL INFRASTRUCTURE CASE STUDY

The critical infrastructure understudy corresponds to the case study of TÜPRAŞ provided, examined in the context of EU funded research project FACTLOG. TÜPRAŞ refinery, located in Izmit, Turkey, produces various petroleum products such as LPG (Liquefied Petroleum Gas), gasoline, diesel, and naphtha. The refinery is composed of multiple units, each serving a specific role in the production process (e.g., production of LPG, production of gasoline, production of diesel, purification of products). FACTLOG project focuses on the LPG purification unit, i.e., on the various processes that need to be applied to turn the LPG production streams to LPG refined streams to meet specific quality specifications.

LPG a mixture of liquefied hydrocarbon gases C3-C4 (propane and butane), is a valuable energy carrier with numerous industry and transportation uses. It is a by-product of many refinery processes, such as Crude Distillation (CDU), Hydrocracking (HYC), Fluid Catalytic Cracking (FCC), and Platformer. After the production process, some impurities

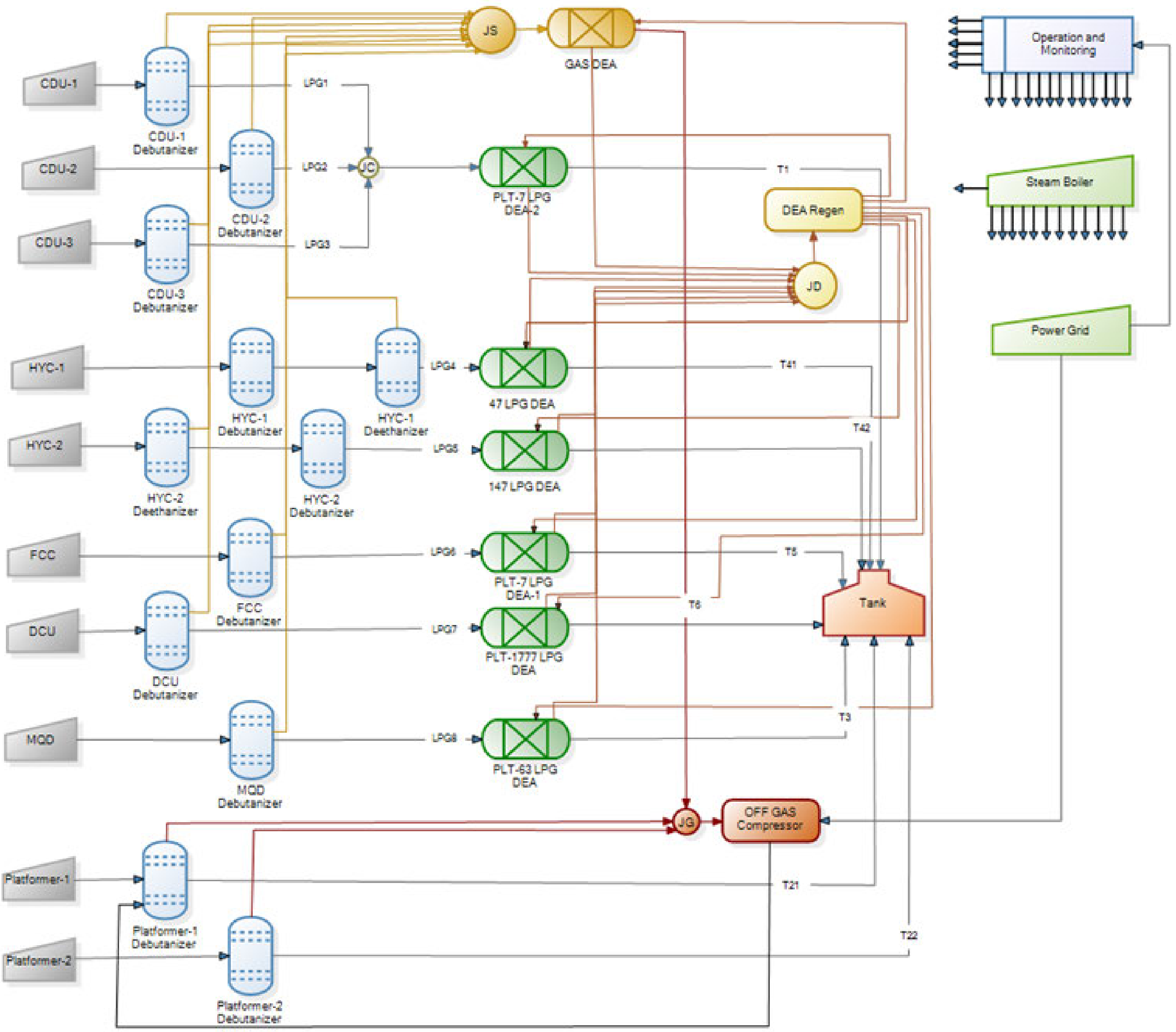


FIGURE 4. Graphical representation of the material flow network. The network models the gas sweetening process for the TUPRAS oil refinery plant.

remain in the LPG that need to be removed (purification). The main purification processes correspond to (i) the removal of Naphtha (C5 and above) in debutanizer columns, (ii) the removal of ethane in Deathanizer columns, and (iii) the removal of sulphur compounds like hydrogen sulphur (H2S) and mercaptan (CH4S) in Amine Absorber Units (AAU).

This study's provided flow network model is representative of any typical LPG purification unit encountered in all oil-refinery industries [89].

Due to permit issues, only the MFN of the purification unit has been provided. Failure probabilities and impact values are assigned based on thorough literature research of similar systems and processes.

Utilizing the provided MFN (see Fig. 4), we identified 22 internal processes, 4 internal junctions, 11 internal inputs, 1 external input, and 1 output node for the part of the production line under study. Moreover, the MFN understudy includes 208 flows of materials and information.

1) PROCESS DEPENDENCY MAPPING

To map the MFN into a dependency graph, we pre-process nodes and flows based on the proposed methodology in Section IV.B. As such, we marked the models' junctions as ignorable. In addition, for each node, we select the required (Primary) resources for nominal operation: for example, for the CDU-1 debutanizer in the MFN C5, S, LPG, C2, Monitor & Control Data, and steam resources are modelled as input flows, but in reality, the debutanizer column requires only LPG gas, steam, and Monitor & Control Data to operate.

Following the pre-processing stage, we map 22 internal processes, four internal inputs, one external input, and one output for the part of the production line under study. Also,

TABLE 3. Mapped MFN nodes and node IDs' association with their respective impact and failure probability values.

Name	ID	Impact	Failure Probability	Source
CDU-1 Debutanizer	P1	4	0.0033	Table 5 & Table 6 [93]
CDU-2 Debutanizer	P2	4	0.0033	
CDU-3 Debutanizer	P3	4	0.0033	
HYC-1 Debutanizer	P4	4	0.0033	
HYC-2 Deethanizer	P5	3	0.0085	Table 13 [94]
FCC Debutanizer	P6	3	0.003	Table 5 & Table 6 [93]
DCU Debutanizer	P7	3	0.003	
MQD Debutanizer	P8	3	0.003	
Platformer-1 Debutanizer	P9	4	0.0033	
Platformer-2 Debutanizer	P10	4	0.0033	
PLT-7 LPG DEA-2	P11	3	0.002147	Table 1 & Table 3 [90]
47 LPG DEA	P12	4	0.002141	
PLT-7 LPG DEA-1	P13	3	0.002147	
HYC-1 Deethanizer	P14	3	0.0085	Table 13 [94]
HYC-2 Debutanizer	P15	4	0.0033	Table 5 & Table 6 [93]
147 LPG DEA	P16	3	0.002147	Table 1 & Table 3 [90]
PLT-1777 LPG DEA	P17	3	0.002147	
PLT-63 LPG DEA	P18	4	0.002141	
GAS DEA	P19	4	0.002141	Table 2 [100] & Table 3 [90]
OFF GAS Compressor	P20	4	0.0025	Table 5 & 6 [93] Table 5 [101]
DEA Regen	P21	2	0.00355	Table 2 & 4 [100]
Operation and Monitoring	P22	4	0.25	Table 2 [104] & Table 5 [105]
CDU-1	I1	3	0.0034	Table 8 [96] & Figure 3 [98]
CDU-2	I2	3	0.0034	
CDU-3	I3	3	0.0034	
HYC-1	I4	4	0.0018	Table 7 [106] Table 5 & 6 [93]
HYC-2	I5	4	0.0018	
FCC	I6	3	0.0075	Table 7 [91] -Table 5 & 6 [93]
DCU	I7	3	0.0034	Table 8 [96] & [98]
MQD	I8	3	0.0034	Figure 7 [99] & [98]
Platformer-1	I9	4	0.004	
Platformer-2	I10	4	0.004	
Steam Boiler	I11	2	0.006	Table 5 & 6 [93] & [107]
Power Grid	I12	4	0.00002	Table 10 [97] & A1–A6 [95]
Tank	O1	5	0.0081	Table 5 & A3 [92]

by marking the required resource for each input, we reduce the total number of flows for mapping to 90. In Table 3, we list the flow network nodes. Flow network nodes depicted use generic terms and IDs in the examples below. Also, in Table 7 given in the Appendix, we list the characterized flows (regular or backup) along with their respective resources, as described in Section IV.B.

The tool automatically maps the material flow network into a risk dependency graph (Fig. 5). Each material flow network node and its respective input and output flows are used to model the asset dependency graph. We should note here that in this specific model, each resource flow corresponds to one dependency.

2) PROCESS DEPENDENCY RISK ANALYSIS

To assess the flow network model, we assign failure probabilities and impact values for each node based on a detailed literature review [90]–[101]. To detect the severity of the consequences for dangerous situations and potential accidents for the physical processes, we searched for previous assessments such as HAZOP, FMEA, and PHA analysis on oil refineries and, in particular, on gas sweetening processes

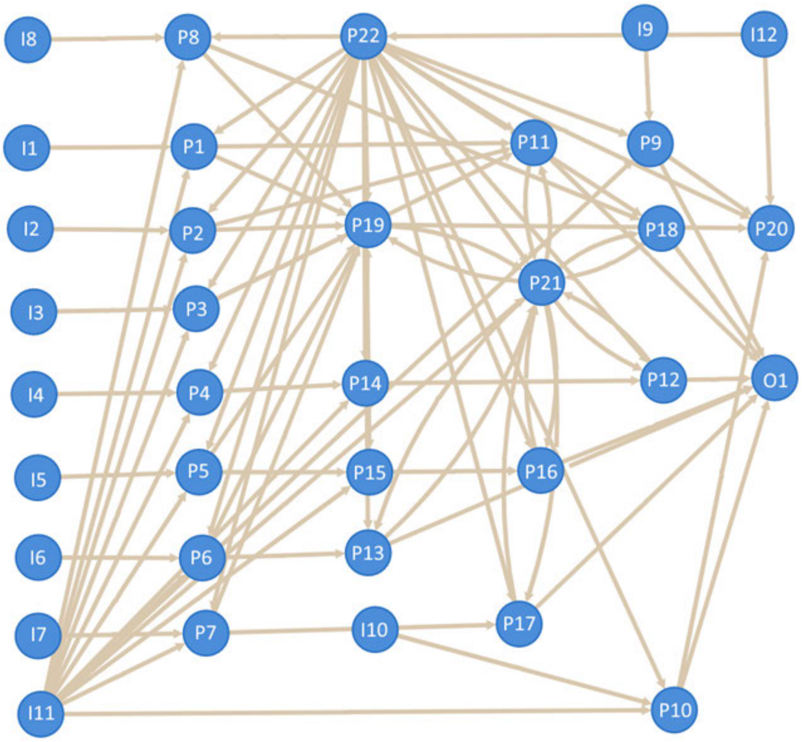


FIGURE 5. Tool graphical representation of the produced dependency graph. Nodes I11 and P22 have the largest outdegree among the modelled dependency graph.

and relevant equipment. Similarly, to identify the probability of a failure mode for a specific process and related equipment, we examined previous FMEA and FTA analyses on oil refineries and gas treatment plants.

A special mention must be made of the monitoring and control process. For this specific process and without having additional information, we consider a typical SCADA network including a control and monitor server placed at the control center, communication links in a corporate network with internet access, and one or more topographically dispersed field sites in the industrial plant containing field devices (i.e., sensors, PLCs, Remote Terminal Unit (RTU) or Intelligent Electronic Device (IED). We estimate the failure probability and the impact for the monitor and control process based on median taken from historical data considering various types of attacks such as (i) replay attacks, (ii) spoofing, (iii) denial of service, (iv) control message modification, (v) unauthorized Write to MTU or RTU, and (vi) RTU response alteration based on literature review [102]–[105].

Table 3 lists the suggested values for both failure probability and impact assigned to each node, respectively. The proposed methodology for the risk analysis allows for engineers and security experts to modify or update probabilities and impact values based on historical data and after consultations with plant operators, engineers, supervisors, and security experts. Specifically, if an organization installs a new security system or changes an industrial machine, experts can decide which attacks or hazards are less likely to cause a particular type of damage by updating the values for the probabilities presented in Table 3.

Based on the assigned failure probabilities and the mapped flows, the tool calculates the likelihood of disruption and, by utilizing the assigned impact values, calculates the risk value of each node in the dependency graph (see Table 4) based on the methods proposed in Section IV.C.1. For example, *GAS DEA (P19),* which models the gas treatment process that removes H2S or other sulfur compounds from the supplied gas using amines (DEA), introduces the *highest risk* with a value of *1.1225* to the dependency graph. On the other hand, as to be expected, the *electricity grid (I12)* introduces the *minimum risk* with a value of *8.00E-05*.

Finally, the tool computed the complete set of risk paths on the risk dependency graph. Paths have an order not greater than 6 (Table 5). In this case, depicted paths correspond to flows from different processes inside the industry. The list below depicts an overview of the computed risk dependency paths.

Thirty-five network flow nodes produced *4033 dependency chains* with orders ranging from three to six and potential *risk values between 0.01 and 1.39*. The tool also computed *the graph's overall dependency risk* with a value of *3637.08* and *the average cumulative dependency* risk with a value of *0.91*. As a result, system engineers and security experts can use this step's output to identify dependency paths with risk values above a threshold value. Also, they can utilize

TABLE 4. Calculated disruption probability and likelihood for each MFN node. Bold letters highlight minimum and maximum likelihood and risk values, respectively.

ID	Likelihood	Risk
P14	0.2633	0.7898
P15	0.2633	1.0531
P16	0.2567	0.7702
P17	0.2565	0.7695
P18	0.2565	1.0260
P19	**0.2806**	**1.1225**
P20	0.2584	1.0336
P21	0.2682	0.5365
P22	0.2500	1.0001
I12	**2.00E-05**	**8.00E-05**
I11	0.0060	0.0120
P1	0.2595	1.0379
P11	0.2616	0.7849
P2	0.2595	1.0379
P3	0.2595	1.0379
O1	0.0272	0.1362
P12	0.2606	1.0424
P13	0.2565	0.7695
I1	0.0034	0.0102
I2	0.0034	0.0102
I3	0.0034	0.0102
I4	0.0018	0.0072
P4	0.2583	1.0332
I5	0.0018	0.0072
P5	0.2622	0.7865
I6	0.0075	0.0225
P6	0.2623	0.7869
I7	0.0034	0.0102
P7	0.2593	0.7778
I8	0.0034	0.0102
P8	0.2593	0.7778
I9	0.0040	0.0160
P9	0.2584	1.0338
I10	0.0040	0.0160
P10	0.2599	1.0397
-		

TABLE 5. Overview of computed dependency paths output from the risk analysis step. Similar paths represent paths with slight variations in sequence and node appearance with the same risk.

Dependency Paths	Risk	# Of Similar Paths
P3 → P19 → P21 → P12 → P21	1.39	123
P3 → P19 → P21 → P16 → P21	1.38	23
P19 → P21 → P12 → P21 → P16 → P21	1.37	34
P19 → P21 → P18 → P21 → P16 → P21	1.36	18
P22 → P1 → P19 → P21 → P16 → P21	1.35	86
P22 → P19 → P21 → P12 → P21	1.34	81
P22 → P15 → P16 → P21 → P19 → P21	1.33	44
P22 → P1 → P11 → P21 → P19 → P21	1.32	88

the number of produced dependency chains, the number of node appearances in paths with risk above a threshold, and the graph's overall dependency risk as metrics to assess and compare the initial system with future redesigns.

In our case, path P3 → P19 → P21 → P12 → P21 is the worst dependency with a risk value of 1.39. Moreover, the tool produced 123 similar paths as the worst dependency with slight variations in sequence and node appearance and, more importantly, the same risk. Hence, based on our analysis, nodes P21, P19, and P22 are the most critical. In particular, nodes P21 and P19 have the highest frequency of occurrence on paths with risk values from 1.39 to 1.36, and node P22 follows in frequency occurrence nodes P19, P21 on dependency paths with risk values from 1.35 to 1.32. *Therefore, Anime (DEA) regeneration (P21) is deemed the most critical process, followed by the gas treatment process (P19) and the operation and monitoring process (P22)*. That is to be expected as the Anime (DEA) regeneration process supplies multiple processes with DEA and at the same time receives DEA from multiple processes for treatment, creating multiple cycles of dependencies, thus increasing the risk of disruption due to an attack or a hazard realized.

3) CRITICAL FLOW ANALYSIS

The target is to identify and prioritize critical flows for risk mitigation. To this end, we produce all alternative graphs with minimum risk (MSTs) based on the method discussed in Section IV.E, using the process dependency graph, and we compute the removal rates for the removed high-risk dependencies. The tool produced *60 alternate MST* with cumulative *risk values between 5.8008 and 5.8015*; each one has a total number of 34 dependencies. For comparison, the initial dependency graph has a summative risk value of 57.90 and total dependencies of 89. We should note here that the produced MST are not applicable redesigns or mitigation measures. *Compared to the initial dependency graph, the produced MSTs have minimized risk, but the removed dependencies represent required material or information flows for nominal system operation*. Nevertheless, the removed dependencies produced during the MST production can indicate where engineers and security experts should focus their efforts to minimize the overall system risk. Table 6 lists a set of high-risk dependencies scored based on how many have been removed from the total number of produced MSTs.

The tool outputs a total number of 109 high-risk dependencies with scores from 40 to 60 (Table 6). System engineers and security experts can use this step's output (Removed Dependencies Report) to identify dependencies, thus system flows, with scores and frequency of appearance above a threshold value.

Our analysis on the remove dependency report shows that dependencies from node P22 appear with higher frequency

TABLE 6. **Removed Dependencies Report. The report presents a set of removed high-risk dependencies product of alternate MST production. The score value (removal rate) depicts the number of MSTs from which the particular dependency has been removed.**

Dependency		Risk	Score	Dependency		Risk	Score
Source	Target			Source	Target		
P12	O1	1.0424	60	P1	P19	1.0379	60
P9	O1	1.0338	60	**P22**	**P17**	1.0001	57
P8	P19	0.7778	60	P7	P17	0.7778	60
P22	**P9**	1.0001	57	**P22**	**P16**	1.0001	57
P22	**P12**	1.0001	57	P15	P16	1.0531	60
P11	P21	0.7849	60	P4	P14	1.0332	60
P22	P8	1.0001	57	P12	P21	1.0424	60
P8	P18	0.7778	60	P19	P20	1.1225	60
P22	**P6**	1.0001	57	**P22**	**P21**	1.0001	57
P14	P12	0.7898	60	**P22**	**P7**	1.0001	57
P18	O1	1.0260	60	P6	P19	0.7869	60
P2	P19	1.0379	60	P14	P19	0.7898	60
P13	O1	0.7695	40	**P22**	**P14**	1.0001	57
P22	P5	1.0001	57	**P22**	**P11**	1.0001	57
P17	P21	0.7695	60	**P22**	**P15**	1.0001	57
P22	**P4**	1.0001	57	P1	P11	1.0379	60
P3	P11	1.0379	60	**P22**	**P18**	1.0001	57
P19	P21	1.1225	60	P2	P11	1.0379	60
P18	P21	1.0260	60	P3	P19	1.0379	60
P22	**P3**	1.0001	57	**P22**	**P19**	1.0001	57
P22	P1	1.0001	57	P5	P19	0.7865	60
P6	P13	0.7869	60	**P22**	**P20**	1.0001	60
P13	P21	0.7695	60	P10	P20	1.0397	60
P22	**P2**	1.0001	57	P16	O1	0.7702	40
P11	O1	0.7849	60	P9	P20	1.0338	60
P17	O1	0.7695	40	**P22**	**P10**	1.0001	57
P22	**P13**	1.0001	57	P7	P19	0.7778	60
P10	O1	1.0397	60	P16	P21	0.7702	60
P5	P15	0.7865	60			-	

(21 times) than all the other source nodes (2 times). *That dictates the importance of the information flows from the operation and monitoring process (P22)*. That is to be expected based on the high risk the specific process introduces to the system, and it is in line with the results from our dependency analysis.

4) PROCESS INFLUENCE ANALYSIS

To identify and prioritize critical flow network nodes based on their influence on the modelled dependency graph, we precompute the centrality metric values for each node using the process dependency graph, as described in Section IV.F, and highlight the maximum values. The results of both the closeness and the eigenvector centrality metrics are presented in Table 7. Based on the results, node P21 presents the highest closeness centrality while node I4 presents the minimum. On the contrary, node P22 presents the highest eigenvector centrality while node I4 presents the minimum. *The anime (DEA) regeneration process (P21) is characterized as a critical node as it is in a position, with its relationships (output flows), to spread risk quickly and to a large portion of the system*. The operation & monitoring process is deemed the most influential because it connects to more influential nodes. *Therefore, a direct attack on the operation & monitoring process (P22) can disrupt critical (strong influence) nodes to the system operation*. The hydrocracking process is the least critical node (low influence), considering both the closeness and the eigenvector centrality metric.

System engineers and security experts can use this step's output to identify a set of critical system components with centrality values above a threshold value. Consequently, based on our centrality analysis, *anime (DEA) regeneration process (P21) and operation & monitoring process (P22) are considered critical nodes with high priority for mitigation measures contrary to the hydrocracking process*.

C. DISCUSSION OF RESULTS

Based on our overall analysis, anime (DEA) regeneration (P21) is deemed the most critical process, followed by the

TABLE 7. Dependency graph closeness and eigen vector centralities. Bold letters highlight minimum and maximum values.

Node	Closeness	Eigen Vector	Node	Closeness	Eigen Vector
	Centrality			Centrality	
P14	0.0139	0.1515	I1	0.0095	0.0165
P15	0.0137	0.1182	I2	0.0095	0.0165
P16	0.0135	0.1838	I3	0.0095	0.0165
P17	0.0137	0.1862	I4	0.0093	0.0115
P18	0.0137	0.1862	P4	0.0135	0.0995
P19	0.0164	0.2953	I5	0.0094	0.0151
P20	0.0141	0.1228	P5	0.0137	0.1328
P21	0.0145	**0.4331**	I6	0.0095	0.0161
P22	**0.0217**	0.4298	P6	0.0139	0.1419
I12	0.0128	0.0662	I7	0.0095	0.0161
I11	0.0179	0.2316	P7	0.0139	0.1419
P1	0.0139	0.1462	I8	0.0095	0.0161
P11	0.0145	0.2198	P8	0.0139	0.1419
P2	0.0139	0.1462	I9	0.0095	0.0134
P3	0.0139	0.1462	P9	0.0139	0.1162
O1	0.0135	0.1666	I10	0.0095	0.0134
P12	0.0135	0.1875	P10	0.0139	0.1162
P13	0.0137	0.1862		-	

gas treatment process (P19) and the operation & monitoring process (P22). Additionally, the removed dependency report dictates a set of flows for risk mitigation. To that end, we discover that information flows from the operation & monitoring process (P22) while not having the highest score have a higher frequency of removal, thus inducing high risk to the system. That is confirmed by our centrality analysis, where the operation & monitoring process is deemed the most influential based on the premise that it connects to more influential nodes. Similarly, the anime (DEA) regeneration (P21) process is a high influence node to spread risk quickly and to a large portion of the system.

Anime (DEA) regeneration process supplies multiple processes (absorbers) with DEA and at the same time receives DEA from them for treatment creating multiple cycles of dependencies. From an engineering perspective, the anime regeneration units have high importance in the chemical absorption process [108]. The gas treatment process (P19) represents an absorber which is undoubtedly the single most crucial operation of gas purification processes [109]. Furthermore, in our case, the gas treatment process receives flows from multiple processes, and it is involved in a circular relation with the Anime (DEA) regeneration process, emerging its criticality.

From a cybersecurity perspective, the operation and control processes and systems are crucial, as they are more open and more vulnerable to cyber-attacks due to existing vulnerabilities [32], [110]. We observed that the monitor and control process, while not part of the highest risk dependency path, has been identified as a critical node by our tool. That is to be expected and considered valid because we are not interested only in cybersecurity threats and vulnerabilities but also hazards and safety issues that initiate from physical processes.

VI. CONCLUSION

The proposed framework following the standard development system lifecycle incorporates security and safety criteria into the design process phase. To achieve that, it provides a clear roadmap for engineers and security experts to identify security and safety risks early in the design phase and address them accordingly, considering the system's quantitative and qualitative requirements.

The proposed methodology can model a CI's production process underlying components and the cyber and physical interactions between them as a material flow network providing a holistic view of the system and a better understanding of the dependencies between the production chain's cyber-physical elements.

Our methodology and developed tool can assess the risk of disruptions due to accidental or intentional events and produce weighted risk dependency graphs, presenting how a disruption in one component may affect other dependent components. Producing the MST of a dependency graph depicts the potential of reducing the network's risk. To that end, by utilizing the removed dependency report produced during the MST production, we can identify and prioritize critical flows. Moreover, by utilizing centrality metrics, we can identify critical flow network components prioritized based on their influence. The prioritization assists engineers and security experts in where they should focus their efforts in order to minimize the overall network risk during the design stage.

TABLE 8. Mapped MFN primary flows and associate resources. Flows are characterized as backup based on design decisions.

Source Node	Target Node	Resource	Backup Flow	Source Node	Target Node	Resource	Backup Flow
P1	P11	LPG	No	P14	P19	LPG	No
P2				**P20**	**P9**		**Yes**
P3				P21	P19	DEA	No
P11	O1			P21	P11		
P12				P21	P12		
P13				P21	P16		
I1	P1			P21	P13		
I2	P2			P21	P17		
I3	P3			P21	P18		
I4	P4			P22	P19	Monitor & Control Data	
I5	P5			P22	P1		
I6	P6			P22	P2		
I7	P7			P22	P3		
I8	P8			P22	P4		
I9	P9			P22	P5		
I10	P10			P22	P14		
P4	P14			P22	P15		
P5	P15			P22	P6		
P14	P12			P22	P11		
P15	P16			P22	P12		
P16	O1			P22	P16		
P7	P17			P22	P13		
P6	P13			P22	P17		
P17	O1			P22	P7		
P8	P18			P22	P21		
P18	O1			I12	P20	Electricity	
P9				I11	P1	Steam	
P10				I11	P2		
P1	P19			I11	P3		
P2				I11	P4		
P3				I11	P5		
P5				I11	P15		
P6				I11	P14		
P7				I11	P6		
P8				I11	P7		
P19	P20			I11	P8		
P9				I11	P9		
P10				I11	P10		
P19	P21			P22		Monitor & Control Data	
P11				P22	P9		
P12				P22	P8		
P16				P22	P18		
P13				P22	P20		
P17				I11	P21	Steam	
P18				I12	P22	Electricity	

The evaluation results from the pilot study in a part of the production line of an existing critical industrial infrastructure show that the presented approach is effective and trustworthy. To that end, our approach supports the proactive study of critical industrial infrastructures with large-scale production chain dependency scenarios advancing the concept of security and safety by design in critical infrastructure protection.

A. RESTRICTIONS AND FUTURE WORK

The presented approach has certain limitations. Like other empirical risk approaches that analyse dependencies, it relies on previous security and safety assessments on related industries and physical components to evaluate the impact and estimate failure probabilities. Our analysis highly depends on the level of detail and quality of the modelled system. By utilizing MFNs, we create and assess a model of reality, and as such, results may miss actual risks. Our framework, especially during mitigation control selection, depends on the subjective opinion of the decision-maker as the tool cannot "guess" erroneous or leftover risks on a design. Also, in our approach, we consider all the required input resources of the modelled flow network nodes as equally important, although in reality, some resources may be more important than others. For modelling disruption of flows, we utilize a binary approach to address availability, but that is not always the case for CI in the industrial sector. Attacks may modify the input quantity of a resource and alter the quality of the output product. These kinds of attacks are not immediately noticed and threaten the integrity of the provided services. Moreover, process mapping is essentially a base that cannot possibly analyse and describe cross-process risks, such as a 3rd-party data monitoring company having a DoS that in turn affects data aggregation from the historian, thus losing data and potentially harming a process in the mid-term future when trying to adjust/optimize it. Future work should concentrate on overcoming the limitations mentioned above.

APPENDIX

See Table 8.

ACKNOWLEDGMENT

The results presented in this paper reflect only the authors' views and the EU is not liable for any use that may be made of the information contained therein. The research aimed at developing innovative methodologies and technologies for implementing the National Cybersecurity Strategy of Greece (2020–2025).

REFERENCES

[1] S. U. Rehman, M. Ceglia, S. Siddiqui, and V. Gruhn, "Towards importance of security for cyber-physical systems/Internet-of-Things," in *Proc. 8th Int. Conf. Softw. Inf. Eng.*, Cairo Egypt, Apr. 2019, pp. 151–155, doi: 10.1145/3328833.3328855.

[2] C. W. Axelrod, "Managing the risks of cyber-physical systems," in *Proc. IEEE Long Island Syst., Appl. Technol. Conf. (LISAT)*, Farmingdale, NY, USA, May 2013, pp. 1–6, doi: 10.1109/LISAT.2013.6578215.

[3] S. Karnouskos, "Stuxnet worm impact on industrial cyber-physical system security," in *Proc. 37th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Melbourne, VIC, Australia, Nov. 2011, pp. 4490–4494, doi: 10.1109/IECON.2011.6120048.

[4] J. Slay and M. Miller, "Lessons learned from the Maroochy water breach," in *Critical Infrastructure Protection*. Boston, MA, USA: Springer, 2008, pp. 73–82.

[5] G. Bakirtzis, T. Sherburne, S. Adams, B. M. Horowitz, P. A. Beling, and C. H. Fleming, "An ontological metamodel for cyber-physical system safety, security, and resilience coengineering," *Softw. Syst. Model.*, to be published, doi: 10.1007/s10270-021-00892-z.

[6] N. Leveson and J. Thomas, *STPA Handbook*. 2018. [Online]. Available: https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf

[7] A. Riel, C. Kreiner, R. Messnarz, and A. Much, "An architectural approach to the integration of safety and security requirements in smart products and systems design," *CIRP Ann.*, vol. 67, no. 1, pp. 173–176, 2018, doi: 10.1016/j.cirp.2018.04.022.

[8] R. Setola, E. Luiijf, and M. Theocharidou, "Critical infrastructures, protection and resilience," in *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach*, R. Setola, V. Rosato, E. Kyriakides, and E. Rome, Eds. Cham, Switzerland: Springer, 2016, pp. 1–18, doi: 10.1007/978-3-319-51043-9_1.

[9] Evaluation Study of Council Directive 2008/114. (Jan. 15, 2020). *Evaluation Study of Council Directive 2008/114 on the Identification and Designation of European Critical Infrastructures and the Assessment of the Need to Improve Their Protection*. EU Publications, Accessed: Feb. 14, 2021. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/e813196c-b041-11ea-bb7a-01aa75ed71a1/language-en

[10] Presidential Policy Directive 21, *Critical Infrastructure Security and Resilience*, 2013.

[11] C. Eckert and O. Isaksson, "Safety margins and design margins: A differentiation between interconnected concepts," *Proc. CIRP*, vol. 60, pp. 267–272, Jan. 2017, doi: 10.1016/j.procir.2017.03.140.

[12] D. Hulse, C. Hoyle, K. Goebel, and I. Y. Tumer, "Quantifying the resilience-informed scenario cost sum: A value-driven design approach for functional hazard assessment," *J. Mech. Des.*, vol. 141, no. 2, Feb. 2019, Art. no. 021403, doi: 10.1115/1.4041571.

[13] P. Dedousis, G. Stergiopoulos, G. Arampatzis, and D. Gritzalis, "Towards integrating security in industrial engineering design practices," in *Proc. 18th Int. Conf. Secur. Cryptogr.*, 2021, pp. 161–172, doi: 10.5220/0010544001610172.

[14] H. Lambrecht and M. Schmidt, "Material flow networks as a means of optimizing production systems," *Chem. Eng. Technol.*, to be published, doi: 10.1002/ceat.200900446.

[15] B. Page, V. Wohlgemuth, and M. Raspe, "Material flow analysis for eco-efficiency with material flow network reference models–concepts and case study," Tech. Rep., 2008.

[16] G. Arampatzis, A. Angelis-Dimakis, M. Blind, and D. Assimacopoulos, "A web-based toolbox to support the systemic eco-efficiency assessment in water use systems," *J. Cleaner Prod.*, vol. 138, pp. 181–194, Dec. 2016, doi: 10.1016/j.jclepro.2016.02.065.

[17] T. Funke and T. Becker, "Complex networks of material flow in manufacturing and logistics: Modeling, analysis, and prediction using stochastic block models," *J. Manuf. Syst.*, vol. 56, pp. 296–311, Jul. 2020, doi: 10.1016/j.jmsy.2020.06.015.

[18] B. Page and V. Wohlgemuth, "Advances in environmental informatics: Integration of discrete event simulation methodology with ecological material flow analysis for modelling eco-efficient systems," *Proc. Environ. Sci.*, vol. 2, pp. 696–705, Jan. 2010, doi: 10.1016/j.proenv.2010.10.079.

[19] V. Wohlgemuth, B. Page, and W. Kreutzer, "Combining discrete event simulation and material flow analysis in a component-based approach to industrial environmental protection," *Environ. Model. Softw.*, vol. 21, no. 11, pp. 1607–1617, Nov. 2006, doi: 10.1016/j.envsoft.2006.05.015.

[20] (2021). *Tüpraş–Refineries*. Accessed: Sep. 1, 2021. [Online]. Available: https://tupras.com.tr/en/rafineries

[21] (2019). *FACTLOG Project*. CORDIS | European Commission. Accessed: Sep. 1, 2021. [Online]. Available: https://cordis.europa.eu/project/id/869951

[22] X. Lyu, Y. Ding, and S. Yang, "Safety and security risk assessment in cyber-physical systems," *IET Cyber-Phys. Syst.: Theory Appl.*, vol. 4, no. 3, pp. 221–232, Sep. 2019, doi: 10.1049/iet-cps.2018.5068.

[23] U. D. Ani, J. D. McK Watson, J. R. C. Nurse, A. Cook, and C. Maples, "A review of critical infrastructure protection approaches: Improving security through responsiveness to the dynamic modelling landscape," in *Proc. Living Internet Things (IoT)*, London, U.K., 2019, p. 6, doi: 10.1049/cp.2019.0131.

[24] Y. Peng, T. Lu, J. Liu, Y. Gao, X. Guo, and F. Xie, "Cyber-physical system risk assessment," in *Proc. 9th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Beijing, China, Oct. 2013, pp. 442–447, doi: 10.1109/IIH-MSP.2013.116.

[25] R. M. Blank, "Guide for conducting risk assessments," National Institute of Standards and Technology, Gaithersburg, MD, USA, Tech. Rep. NIST SP 800-30r1, 2012, doi: 10.6028/NIST.SP.800-30r1.

[26] *Information Technology-Security Techniques-Information Security Management Systems-Requirements*, Standard BS ISO/IEC 27001, BSI, London, U.K., 2013.

[27] *COBIT 5: A Business Framework for the Governance and Management of Enterprise IT*, ISACA, Schaumburg, IL, USA, 2012.

[28] (Dec. 2016). *The 62443 Series of Standards–Industrial Automation and Control Systems*. Accessed: Apr. 16, 2021. [Online]. Available: https://gca.isa.org/hubfs/ISAGCAQuickStartGuideFINAL.pdf

[29] R. S. H. Piggin, "Development of industrial cyber security standards: IEC 62443 for scada and industrial control system security," in *Proc. IET Conf. Control Autom.: Uniting Problems Solutions*, Birmingham, U.K., 2013, p. 11, doi: 10.1049/cp.2013.0001.

[30] D. Gritzalis, G. Stergiopoulos, V. Kouktzoglou, and M. Theocharidou, "A process-based dependency risk analysis methodology for critical infrastructures," *Int. J. Crit. Infrastruct.*, vol. 13, nos. 2–3, p. 184, 2017, doi: 10.1504/IJCIS.2017.10009258.

[31] Y. Cherdantseva, P. Burnap, A. Blyth, P. Eden, K. Jones, H. Soulsby, and K. Stoddart, "A review of cyber security risk assessment methods for SCADA systems," *Comput. Secur.*, vol. 56, pp. 1–27, Feb. 2016, doi: 10.1016/j.cose.2015.09.009.

[32] E. Johansson, T. Sommestad, and M. Ekstedt, "Issues of cyber security in SCADA-systems–on the importance of awareness," in *Proc. IET Conf. Publications*, Prague, Czech Republic, 2009, p. 969, doi: 10.1049/cp.2009.1099.

[33] Z. Basnight, J. Butts, J. Lopez, and T. Dube, "Firmware modification attacks on programmable logic controllers," *Int. J. Crit. Infrastruct. Protection*, vol. 6, no. 2, pp. 76–84, Jun. 2013, doi: 10.1016/j.ijcip.2013.04.004.

[34] H. Yoo and I. Ahmed, "Control logic injection attacks on industrial control systems," in *ICT Systems Security and Privacy Protection*. Cham, Switzerland: Springer, 2019, pp. 33–48.

[35] A. Roy, D. S. Kim, and K. S. Trivedi, "Scalable optimal countermeasure selection using implicit enumeration on attack countermeasure trees," in *Proc. IEEE/IFIP Int. Conf. Dependable Syst. Netw. (DSN)*, Boston, MA, USA, Jun. 2012, pp. 1–12, doi: 10.1109/DSN.2012.6263940.

[36] A. Cavoukian and M. Dixon, *Privacy and Security by Design: An Enterprise Architecture Approach*, Information and Privacy Commissioner of Ontario, Toronto, ON, Canada, 2013.

[37] B. Filkins. (Feb. 2020). *Security by Design: A Systems Road Map Ap-Proach*. SANS. Accessed: Feb. 13, 2021. [Online]. Available: https://www.sans.org/reading-room/whitepapers/awareness/security-design-systems-road-map-approach-39370

[38] T. Fagade, K. Maraslis, and T. Tryfonas, "Towards effective cybersecurity resource allocation: The Monte Carlo predictive modelling approach," *Int. J. Crit. Infrastruct.*, vol. 13, nos. 2–3, p. 152, 2017, doi: 10.1504/IJCIS.2017.088235.

[39] B. Johnson, V. Chalishazar, E. Cotilla-Sanchez, and T. K. A. Brekken, "A Monte Carlo methodology for earthquake impact analysis on the electrical grid," *Electr. Power Syst. Res.*, vol. 184, Jul. 2020, Art. no. 106332, doi: 10.1016/j.epsr.2020.106332.

[40] E. Ferrario, N. Pedroni, and E. Zio, "Evaluation of the robustness of critical infrastructures by hierarchical graph representation, clustering and Monte Carlo simulation," *Rel. Eng. Syst. Saf.*, vol. 155, pp. 78–96, Nov. 2016, doi: 10.1016/j.ress.2016.06.007.

[41] E. O. Ebewele and M. H. Al-Marzouqi, "Regeneration of solvent for CO_2 capture: A review," in *Proc. 6th Int. Conf. Renew. Energy: Gener. Appl. (ICREGA)*, Al Ain, United Arab Emirates, Feb. 2021, pp. 163–167, doi: 10.1109/ICREGA50506.2021.9388298.

[42] M. Waqas, A. S. Aburiazaiza, R. Miandad, M. Rehan, M. A. Barakat, and A. S. Nizami, "Development of biochar as fuel and catalyst in energy recovery technologies," *J. Cleaner Prod.*, vol. 188, pp. 477–488, Jul. 2018, doi: 10.1016/j.jclepro.2018.04.017.

[43] A. Nouri. Gharahasanlou, A. Mokhtarei, A. Khodayarei, and M. Ataei, "Fault tree analysis of failure cause of crushing plant and mixing bed Hall at Khoy cement factory in Iran," *Case Stud. Eng. Failure Anal.*, vol. 2, no. 1, pp. 33–38, Apr. 2014, doi: 10.1016/j.csefa.2013.12.006.

[44] H. Watson, *Launch Control Safety Study*, Bell Laboratories, Holmdel, NJ, USA, 1961.

[45] T. A. Carbone and D. D. Tippett, "Project risk management using the project risk FMEA," *Eng. Manage. J.*, vol. 16, no. 4, pp. 28–35, Dec. 2004, doi: 10.1080/10429247.2004.11415263.

[46] H. A. Duckworth and R. A. Moore, *Social Responsibility: Failure Mode Effects and Analysis*. Boca Raton, FL, USA: CRC Press, 2010, doi: 10.1201/EBK1439803721.

[47] T. A. Kletz, "Hazop—Past and future," *Rel. Eng. Syst. Saf.*, vol. 55, no. 3, pp. 263–266, Mar. 1997, doi: 10.1016/S0951-8320(96)00100-7.

[48] H. G. Lawley, "Operability studies and hazard analysis," *Chem. Eng. Prog.*, vol. 70, pp. 105–116, Jan. 1974.

[49] A. Abdulkhaleq, S. Wagner, and N. Leveson, "A comprehensive safety engineering approach for software-intensive systems based on STPA," *Proc. Eng.*, vol. 128, pp. 2–11, Jan. 2015, doi: 10.1016/j.proeng.2015.11.498.

[50] T. Ishimatsu, N. Leveson, J. Thomas, M. Katahira, Y. Miyamoto, and H. Nakao, "Modeling and hazard analysis using STPA," Tech. Rep., 2010.

[51] I. Friedberg, K. McLaughlin, P. Smith, D. Laverty, and S. Sezer, "STPA-SafeSec: Safety and security analysis for cyber-physical systems," *J. Inf. Secur. Appl.*, vol. 34, pp. 183–196, Jun. 2017, doi: 10.1016/j.jisa.2016.05.008.

[52] S. Kriaa, L. Pietre-Cambacedes, M. Bouissou, and Y. Halgand, "A survey of approaches combining safety and security for industrial control systems," *Rel. Eng. Syst. Saf.*, vol. 139, pp. 156–178, Jul. 2015, doi: 10.1016/j.ress.2015.02.008.

[53] J. Axelsson and A. Kobetski, "Towards a risk analysis method for systems-of-systems based on systems thinking," in *Proc. Annu. IEEE Int. Syst. Conf. (SysCon)*, Vancouver, BC, Canada, Apr. 2018, pp. 1–8, doi: 10.1109/SYSCON.2018.8369501.

[54] A. Kobetski and J. Axelsson, "Towards safe and secure systems of systems: Challenges and opportunities," in *Proc. Symp. Appl. Comput.*, Marrakech Morocco, Apr. 2017, pp. 1803–1806, doi: 10.1145/3019612.3028252.

[55] S. Plósz, C. Schmittner, and P. Varga, "Combining safety and security analysis for industrial collaborative automation systems," in *Computer Safety, Reliability, and Security*. Cham, Switzerland: Springer, 2017, pp. 187–198.

[56] A. R. Parkinson, R. Balling, and J. D. Hedengren, *Optimization Methods for Engineering Design*, 2nd ed. Provo, UT, USA: Brigham Young University, 2018. [Online]. Available: https://learn.skillman.eu/pluginfile.php/921/mod_resource/content/0/optimization_book.pdf

[57] S. Ziegler, A. Skarmeta, J. Bernal, E. E. Kim, and S. Bianchi, "ANASTACIA: Advanced networked agents for security and trust assessment in CPS IoT architectures," in *Proc. Global Internet Things Summit (GIoTS)*, Geneva, Switzerland, Jun. 2017, pp. 1–6, doi: 10.1109/GIOTS.2017.8016285.

[58] Y. Chen, L. He, J. Li, and S. Zhang, "Multi-criteria design of shale-gas-water supply chains and production systems towards optimal life cycle economics and greenhouse gas emissions under uncertainty," *Comput. Chem. Eng.*, vol. 109, pp. 216–235, Jan. 2018, doi: 10.1016/j.compchemeng.2017.11.014.

[59] Z. Rizki, A. E. M. Janssen, G. D. H. Claassen, R. M. Boom, and A. van der Padt, "Multi-criteria design of membrane cascades: Selection of configurations and process parameters," *Separat. Purification Technol.*, vol. 237, Apr. 2020, Art. no. 116349, doi: 10.1016/j.seppur.2019.116349.

[60] *Security for Industrial Automation and Control Systems, Part 3-2: Security Risk Assessment for System Design*, International Society of Automation, Standard ANSI/ISA-62443-3-2-2020, 2020.

[61] A. C. Caputo, P. M. Pelagagge, and P. Salini, "AHP-based methodology for selecting safety devices of industrial machinery," *Saf. Sci.*, vol. 53, pp. 202–218, Mar. 2013, doi: 10.1016/j.ssci.2012.10.006.

[62] P. Kotzanikolaou, M. Theoharidou, and D. Gritzalis, "Assessing N-order dependencies between critical infrastructures," *Int. J. Crit. Infrastruct.*, vol. 9, nos. 1–2, p. 93, 2013, doi: 10.1504/IJCIS.2013.051606.

[63] G. Stergiopoulos, P. Kotzanikolaou, M. Theocharidou, G. Lykou, and D. Gritzalis, "Time-based critical infrastructure dependency analysis for large-scale and cross-sectoral failures," *Int. J. Crit. Infrastruct. Protection*, vol. 12, pp. 46–60, Mar. 2016, doi: 10.1016/j.ijcip.2015.12.002.

[64] B. Bush, L. Dauelsberg, R. Leclaire, D. Powell, and S. DeLand, "Critical infrastructure protection decision support system (CIP/DSS) project overview," Tech. Rep., Jan. 2005.

[65] G. Stergiopoulos, P. Dedousis, and D. Gritzalis, "Automatic network restructuring and risk mitigation through business process asset dependency analysis," *Comput. Secur.*, vol. 96, Sep. 2020, Art. no. 101869, doi: 10.1016/j.cose.2020.101869.

[66] P. H. Brunner and H. Rechberger, *Handbook of Material Flow Analysis: For Environmental, Resource, and Waste Engineers*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2020.

[67] A. Tiwana and M. Keil, "The one-minute risk assessment tool," *Commun. ACM*, vol. 47, no. 11, pp. 73–77, Nov. 2004, doi: 10.1145/1029496.1029497.

[68] B. Adenso-Diaz, C. Mena, S. García-Carbajal, and M. Liechty, "The impact of supply network characteristics on reliability," *Supply Chain Manage., Int. J.*, vol. 17, no. 3, pp. 263–276, Apr. 2012, doi: 10.1108/13598541211227108.

[69] L. V. Snyder, Z. Atan, P. Peng, Y. Rong, A. J. Schmitt, and B. Sinsoysal, "OR/MS models for supply chain disruptions: A review," *IIE Trans.*, vol. 48, no. 2, pp. 89–109, 2016, doi: 10.1080/0740817X.2015.1067735.

[70] M. Eid and V. Rosato, "Critical infrastructure disruption scenarios analyses via simulation," in *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach*, R. Setola, V. Rosato, E. Kyriakides, and E. Rome, Eds. Cham, Switzerland: Springer, 2016, pp. 43–61, doi: 10.1007/978-3-319-51043-9_3.

[71] J. Pearl, *Probabilistic Reasoning in Intelligent Systems*. Amsterdam, The Netherlands: Elsevier, 1988, doi: 10.1016/C2009-0-27609-4.

[72] A. Käki, A. Salo, and S. Talluri, "Disruptions in supply networks: A probabilistic risk assessment approach," *J. Bus. Logistics*, vol. 36, no. 3, pp. 273–287, Sep. 2015, doi: 10.1111/jbl.12086.

[73] K. Zhou, A. Martin, and Q. Pan, "The belief noisy-OR model applied to network reliability analysis," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 24, no. 6, pp. 937–960, Dec. 2016, doi: 10.1142/S0218488516500434.

[74] F. M. Akhmedjanov, *Reliability Databases: State-of-the-Art and Perspectives*. 2001.

[75] M. Held and R. M. Karp, "The traveling-salesman problem and minimum spanning trees," *Oper. Res.*, vol. 18, no. 6, pp. 1138–1162, Dec. 1970, doi: 10.1287/opre.18.6.1138.

[76] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. Amer. Math. Soc.*, vol. 7, no. 1, p. 48, Jan. 1956, doi: 10.1090/S0002-9939-1956-0078686-7.

[77] S. Pettie and V. Ramachandran, "An optimal minimum spanning tree algorithm," in *Automata, Languages and Programming*, vol. 1853, U. Montanari, J. D. P. Rolim, and E. Welzl, Eds. Berlin, Germany: Springer, 2000, pp. 49–60, doi: 10.1007/3-540-45022-X_6.

[78] R. C. Prim, "Shortest connection networks and some generalizations," *Bell Syst. Tech. J.*, vol. 36, no. 6, pp. 1389–1401, 1957, doi: 10.1002/j.1538-7305.1957.tb01515.x.

[79] S. Kiesling, J. Klünder, D. Fischer, K. Schneider, and K. Fischbach, "Applying social network analysis and centrality measures to improve information flow analysis," in *Product-Focused Software Process Improvement*, vol. 10027, P. Abrahamsson, A. Jedlitschka, A. Nguyen Duc, M. Felderer, S. Amasaki, and T. Mikkonen, Eds. Cham, Switzerland: Springer, 2016, pp. 379–386, doi: 10.1007/978-3-319-49094-6_25.

[80] L. Maccari, Q. Nguyen, and R. Lo Cigno, "On the computation of centrality metrics for network security in mesh networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6, doi: 10.1109/GLOCOM.2016.7842049.

[81] F. A. Rodrigues, "Network centrality: An introduction," in *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, E. Macau, Ed. Cham, Switzerland: Springer, 2019, pp. 177–196, doi: 10.1007/978-3-319-78512-7_10.

[82] G. Stergiopoulos, P. Kotzanikolaou, M. Theocharidou, and D. Gritzalis, "Risk mitigation strategies for critical infrastructures based on graph centrality analysis," *Int. J. Crit. Infrastruct. Protection*, vol. 10, pp. 34–44, Sep. 2015, doi: 10.1016/j.ijcip.2015.05.003.

[83] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkevičiūtė, C. Suo, and A. Fornito, "Consistency and differences between centrality measures across distinct classes of networks," *PLoS ONE*, vol. 14, no. 7, Jul. 2019, Art. no. e0220061, doi: 10.1371/journal.pone.0220061.

[84] P. Bonacich, "Power and centrality: A family of measures," *Amer. J. Sociol.*, vol. 92, no. 5, pp. 1170–1182, 1987, doi: 10.1086/228631.

[85] A. Bavelas, "Communication patterns in task-oriented groups," *J. Acoust. Soc. Amer.*, vol. 22, no. 6, pp. 725–730, Nov. 1950, doi: 10.1121/1.1906679.

[86] L. C. Freeman, "Centrality in social networks conceptual clarification," *Soc. Netw.*, vol. 1, no. 3, pp. 215–239, Jan. 1978, doi: 10.1016/0378-8733(78)90021-7.

[87] M. Van Eeten, A. Nieuwenhuijs, E. Luiijf, M. Klaver, and E. Cruz, "The state and the threat of cascading failure across critical infrastructures: The implications of empirical evidence from media incident reports," *Public Admin.*, vol. 89, no. 2, pp. 381–400, Jun. 2011, doi: 10.1111/j.1467-9299.2011.01926.x.

[88] (2000). *Neo4J Graph Database*. Neo4j Graph Database Platform. Accessed: Feb. 23, 2021. [Online]. Available: https://neo4j.com/product/neo4j-graph-database/

[89] A. Bahadori, *Natural Gas Processing: Technology and Engineering Design*. Amsterdam, The Netherlands: Elsevier, 2014.

[90] A. Askarian, M. J. Jafari, L. Omidi, M. R. Miri Lavasani, L. Taghavi, and A. Ashori, "Hazard identification and risk assessment in two gas refinery units," *Health Scope*, vol. 7, no. 1, Feb. 2018, Art. no. e68252, doi: 10.5812/jhealthscope.68252.

[91] D. Gritzalis, G. Iseppi, A. Mylonas, and V. Stavrou, "Exiting the risk assessment maze: A meta-survey," *ACM Comput. Surv.*, vol. 51, no. 1, p. 11, Jan. 2018.

[92] J. Fuentes-Bargues, M. González-Cruz, C. González-Gaya, and M. Baixauli-Pérez, "Risk analysis of a fuel storage terminal using HAZOP and FTA," *Int. J. Environ. Res. Public Health*, vol. 14, no. 7, p. 705, Jun. 2017, doi: 10.3390/ijerph14070705.

[93] A. Ibrahim and A. U. El-Nafaty, "Assessment of the reliability of fractionator column of the Kaduna refinery using failure modes effects and criticality analysis (FMECA)," *Amer. J. Eng. Res.*, vol. 5, no. 2, pp. 101–108, 2016.

[94] S. M. Lavasani, A. Zendegani, and M. Celik, "An extension to fuzzy fault tree analysis (FFTA) application in petrochemical process industry," *Process Saf. Environ. Protection*, vol. 93, pp. 75–88, Jan. 2015, doi: 10.1016/j.psep.2014.05.001.

[95] P. K. Marhavilas, M. Filippidis, G. K. Koulinas, and D. E. Koulouriotis, "A HAZOP with MCDM based risk-assessment approach: Focusing on the deviations with economic/health/environmental impacts in a process industry," *Sustainability*, vol. 12, no. 3, p. 993, Jan. 2020, doi: 10.3390/su12030993.

[96] A. Musyafa, R. Kresna, A. A. H. Cordova, and R. D. Noriyati, "HAZOP study and risk assessment in three-phase separator oil and gas exploration farm–East Java, Indonesia," *Adv. Natural Appl. Sci.*, vol. 11, no. 3, p. 77, Mar. 2017.

[97] R. Novo, F. Neirotti, L. Visocaro, and A. Colangelo, "Oil and gas treatment plant–risk analysis of the fire-fighting system and of the gas treatment unit," Tech. Rep., 2017, doi: 10.13140/RG.2.2.27606.75841.

[98] B. Riad, B. Hamid, R. Hind, and Z. Youcef, "Design of an integration frame HAZOP-SIL for safety optimization of a fired heater," in *Proc. Int. Conf. Electr. Sci. Technol. Maghreb (CISTEM)*, Algiers, Algeria, Oct. 2018, pp. 1–6, doi: 10.1109/CISTEM.2018.8613375.

[99] S. Werner and W. Fred, "Assessing safety in distillation column using dynamic simulation and failure mode and effect analysis (FMEA)," *J. Appl. Sci.*, vol. 7, no. 15, pp. 2033–2039, Jul. 2007, doi: 10.3923/jas.2007.2033.2039.

[100] C. You and J. Kim, "Quantitative risk assessment of an amine-based CO_2 capture process," *Korean J. Chem. Eng.*, vol. 37, no. 10, pp. 1649–1659, Oct. 2020, doi: 10.1007/s11814-020-0567-5.

[101] Y. Zennir and R. Bendib, "The dependability control analysis: Applied to centrifugal pumps in a oil petrochemical plant," in *Proc. Int. Conf. Ind. Eng. Syst. Manage. (IESM)*, Seville, Spain, Oct. 2015, pp. 1004–1011, doi: 10.1109/IESM.2015.7380277.

[102] G. Hamoud, R.-L. Chen, and I. Bradley, "Risk assessment of power systems SCADA," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, Toronto, ON, Canada, Jul. 2003, pp. 758–764, doi: 10.1109/PES.2003.1270402.

[103] S. C. Patel, J. H. Graham, and P. A. S. Ralston, "Quantitatively assessing the vulnerability of critical information systems: A new method for evaluating security enhancements," *Int. J. Inf. Manage.*, vol. 28, no. 6, pp. 483–491, Dec. 2008, doi: 10.1016/j.ijinfomgt.2008.01.009.

[104] S. Patel and J. Zaveri, "A risk-assessment model for cyber attacks on information systems," *J. Comput.*, vol. 5, no. 3, pp. 352–359, Mar. 2010, doi: 10.4304/jcp.5.3.352-359.

[105] D. Upadhyay and S. Sampalli, "SCADA (Supervisory control and data Acquisition) systems: Vulnerability assessment and security recommendations," *Comput. Secur.*, vol. 89, Feb. 2020, Art. no. 101666, doi: 10.1016/j.cose.2019.101666.

[106] I. Nakamanuruck, V. Rungreunganun, and S. Talabgaew, "Reliability analysis for refinery plants," *Appl. Sci. Eng. Prog.*, vol. 10, no. 1, 2017, doi: 10.14416/j.ijast.2017.01.002.

[107] R. Bendib, Y. Zennir, E.-A. Mechhoud, and S. Bouziane, "Risk assessment for a steam generator (1050 G1) Skikda refinery Algeria, using HAZOP and RQA methods," in *Proc. Int. Conf. Adv. Syst. Emergent Technol. (IC_ASET)*, Hammamet, Tunisia, Mar. 2019, pp. 262–267, doi: 10.1109/ASET.2019.8871025.

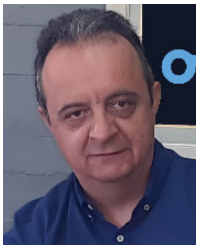
[108] T. Li and T. C. Keener, "A review: Desorption of CO_2 from rich solutions in chemical absorption processes," *Int. J. Greenhouse Gas Control*, vol. 51, pp. 290–304, Aug. 2016, doi: 10.1016/j.ijggc.2016.05.030.

[109] S. Mitra, "A technical report on gas sweetening system," Tech. Rep., 2015, doi: 10.13140/RG.2.1.2061.9360.

[110] B. Miller and D. Rowe, "A survey SCADA of and critical infrastructure incidents," in *Proc. 1st Annu. Conf. Res. Inf. Technol. (RIIT)*, Calgary, AB, Canada, 2012, p. 51, doi: 10.1145/2380790.2380805.

PANAGIOTIS DEDOUSIS received the B.Sc. degree in informatics from the University of Piraeus, Greece, and the M.Sc. degree in information systems from the Athens University of Economics and Business (AUEB), Greece, where he is currently pursuing the Ph.D. degree in information security and critical infrastructure protection. He is also a Researcher with the Information Security & Critical Infrastructure Protection (INFOSEC) Research Group, Department of Informatics, AUEB. His current research interests include information security, critical infrastructure protection, and risk assessment.

GEORGE STERGIOPOULOS received the B.Sc. degree in informatics from the University of Piraeus, Greece, the M.Sc. degree in information systems, and the Ph.D. degree in critical infrastructure protection at software and information interdependency levels from the Athens University of Economics and Business (AUEB), Greece. He is currently an Assistant Professor in cybersecurity with the Department of Information & Communication Systems Engineering, University of the Aegean, Greece. He has been the Technical Director or a Principal Investigator with the INFOSEC Laboratory, AUEB, in numerous funded research projects in the areas of critical infrastructure protection, software security, malware, and network security. He is an Expert in standardization (ISO 27001) and in the European Union's General Data Protection Regulation (GDPR). He has published several papers in peer-reviewed journals and international conferences.



GEORGE ARAMPATZIS received the Diploma and Ph.D. degrees in chemical engineering from the National Technical University of Athens. He is currently an Assistant Professor with the School of Production Engineering and Management (PEM), Technical University of Crete (TUC), and a Visiting Professor with the Mediterranean Agronomic Institute of Chania, Greece. He is the Coordinator of the ''Industrial and Digital Innovations Research Group'' (INDIGO). He has participated in more than 30 research projects at both European and National level as a Technical/a Scientific Coordinator, the Work Package Leader and a Senior Researcher. He is the author of six academic books, three book chapters and of more than 70 scientific publications in international peer-reviewed journals and conference proceedings. His research interests include smart ICT technologies, process, system and service engineering, energy systems management, environmental systems management, water resources management, and decision making.

DIMITRIS GRITZALIS received the B.Sc. degree in mathematics from the University of Patras, Greece, the M.Sc. degree in computer science from the City University of New York, USA, and the Ph.D. degree in information systems security from the University of the Aegean, Greece. He is currently a Professor in cybersecurity with the Department of Informatics, Athens University of Economics and Business (AUEB), Greece. He also works as the Director of the M.Sc. Programme in information systems development & security, and the Director of the INFOSEC Research Group. He has worked as an Associate Rector for Research, an Associate Rector for Financial Affairs, and as the President of the Life-Long Education Center, AUEB. He has also worked as the President of the Greek Computer Society, and as an Associate Data Protection Commissioner of Greece. He has published extensively in peer-reviewed journals and conferences. His current research interests include cybersecurity, risk assessment, critical infrastructure protection, and malware. He serves as an Academic Editor of *Computers & Security*, (Elsevier) and as a Scientific Editor of the *International Journal of Critical Infrastructure Protection*, (Elsevier).