Towards Integrating Security in Industrial Engineering Design Practices

Panagiotis Dedousis[1], George Stergiopoulos[1,2], George Arampatzis[3] and Dimitris Gritzalis[1]
[1]*Dept. of Informatics, Athens University of Economics & Business, Athens, Greece*
[2]*Dept. of Information & Communication Systems Engineering, University of the Aegean, Samos, Greece*
[3]*School of Production Engineering & Management, Technical University of Crete, Chania, Greece*

Keywords: Critical Infrastructure Protection, Component Cascading Failures, Dependency Risk Graphs, Resilience.

Abstract: During the past decades, and especially since the Stuxnet event, there has being a growing concern around the protection of critical infrastructures. Even though the protection of such systems and services has been an international security priority, still, even after all those years, relevant research either focuses on individual ICS systems security (PLC, RTU and SCADA network protection and attacks), or uses high-level models to perform risk assessments, mostly from a system-of-systems scope that studies interdependencies. From an engineering perspective, current approaches address system resilience from an efficiency perspective (i.e. focusing on the availability of physical processes) while neglecting the security dimension of their components. Still, the availability and reliability requirements of such systems are directly affected by security incidents. To our knowledge, there is currently no process to integrate security-by-design in industrial critical infrastructure engineering. To this end, we present a method to integrate security risk assessment analysis into engineering design practices. We do this by modeling internal dependencies between physical components in critical industrial production processes to identify possible hotspots of system failures that are challenging to handle later in the development lifecycle, especially during operation. To validate our approach, we model and assess the present situation in a portion of an actual oil refining plant, thereby establishing a baseline model. Then we introduce risk mitigation measures by altering the design of the baseline model, resulting in a reduction of the overall cascade risk.

1 INTRODUCTION

Critical infrastructures (CIs) provide vital services to the society, depending on the specific sector of their activity, such as energy, water, health, finance, transportation. All sectors serve as the main structure of the economy, security, and health. Their destruction poses one of the greatest threats to physical security, national economic, national public health or safety, or any combination thereof.

For authorities worldwide, minimizing the risk against all potential threats that could damage CIs by enhancing resilience and reliability has become a regulatory requirement (Setola et al., 2016). The US supports this through the NIS Directive (Evaluation study of Council Directive 2008/114, 2020), while the US has published specific frameworks solely to protect CIs. (Parfomak, 2007).

Emerging threats and unconventional attacks on CIs have exposed the limits of traditional risk assessment and risk mitigation efforts (Ani et al., 2019). Some threats cannot be anticipated, and it is not always cost-effective to reduce all potential risks at the minimum level possible.

Even though authorities have long identified the risks behind cyberattacks on CI, still, to our knowledge, there is no work on how to integrate security-by-design principles in industrial critical infrastructure engineering. Investing in system architecture and focusing on resilience during the design stage is much more efficient and less costly than keep funding the protection of systems that are not resilient and secure (Eckert & Isaksson, 2017; Hulse et al., 2019).

In our paper, we present our work towards integrating engineering design practices with security risk assessment and dependency analysis. This integration enables engineers and security experts to identify and

Dedousis, P., Stergiopoulos, G., Arampatzis, G. and Gritzalis, D.
Towards Integrating Security in Industrial Engineering Design Practices.
DOI: 10.5220/0010544001610172
In *Proceedings of the 18th International Conference on Security and Cryptography (SECRYPT 2021)*, pages 161-172
ISBN: 978-989-758-524-1
Copyright © 2021 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

assess security issues at the level of process-based engineered systems (industrial applications). Critical industrial infrastructures are *production systems* with complex *production chains*. Their production chains are networks of *productive activities* and are characterized by *flows* of resources; i.e., physical flows of materials and energy, and flows of monitoring and control information accompanying the physical flows. 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 output.

Our main contribution is a component-centric approach that focuses on each process's input-output resources comprising a system. Engineers are mainly interested in the flows of materials (oil, water, gas) and energy (electrical, chemical, thermal) and not information flows. They also focus only on processes. Instead, we focus on information flows in conjunction with engineering processes in a more holistic way to design such material flows. We utilize material flow analysis (MFA) to model the underline processes and flows of a critical industrial infrastructure production chain into a material flow network graph, thus creating a design model. Also, we utilize a risk assessment and dependency analysis methodology to evaluate the cascade impacts of process disruptions and the overall risk affecting the CI based on the design model. Our approach can identify process points of potential failures in engineering process design. Such points of failure are challenging to handle later in the development lifecycle, especially during operation. By identifying such hotspots, countermeasures can be integrated directly into the design process of the project lifecycle, improving system reliability and resiliency.

We apply our approach to assess a part of the production chain corresponding to Liquefied Petroleum Gas (LPG) purification of an actual oil refining plant under a high-risk scenario. We analyze specific production flows and then implement selected mitigation measures by altering and modifying the baseline network flow design of the LPG model to reduce the overall risk, thus proving/demonstrating the validity of using such measures during the design step of such processes. In summary, we contribute the following:

1. A technique that models material, energy, and information flows using security principles in production chains of critical infrastructures.
2. A modeling approach that maps and converts the assets and the interdependencies of a material flow network into a risk dependency graph based on the existing CI production chain topology.
3. A risk calculation methodology that depicts the probability of a threat to disrupt a CI asset based on a noisy-OR model and a rating scale to evaluate each impact's severity.

The rest of the paper is organized as follows: Section 2 discusses related work and compares critical infrastructure protection methods. Section 3 describes the fundamental building blocks. Section 4 describes step by step the proposed methodology. Section 5 describes the implemented tool and analyses the implementation of the methodology in a real-world example. The conclusion discusses paper results and potential future research.

2 RELATED WORK

Various methodologies are used to determine resilience and evaluate the different dimensions involved in the factors that affect CIs. The methods proposed in the scientific literature mostly focus on: (i) risk assessment methodologies for assessing the risk of CI, (ii) CI interdependencies, (iii) ICS systems security, and (iv) security-by-design approaches. The main intention of such high-level methodologies is to analyze the multi-dimensional impacts of disruptive incidents involving CIs in multiple sectors (Ani et al., 2019).

Traditional risk assessment methodologies usually focus on individual vulnerabilities on already operational systems (BS ISO/IEC 27001, 2013; ISACA. 2012b, 2012; NIST SP 800-30, 2012). However, traditional Risk assessments performed on already established and functioning systems mostly result in added layers of cybersecurity on top of existing systems. Similar approaches adopt the concept of system-of-systems to address safety effectively, security, reliability, and robustness in CIs (Axelsson & Kobetski, 2018; Kobetski & Axelsson, 2017). Although system-of-systems analysis provides a comprehensive top-down overview of the environment in which the CI operates and how risk propagates to and from the system, it is fraught with uncertainty about how constituent systems operate and function.

CI are complex systems that depend on other CI or/and suppliers for their operations. Several authors stressed the importance of modeling CI as interconnected systems to explain the cascade effects due to their strong interdependencies (Min et al., 2007; Rinaldi et al., 2001). Many approaches utilize graph visualization or cascade diagrams to model the interdependencies among CI and assess the cascading risk considering cross-sectoral and cross-border interdependencies (Azzini et al., 2018; Ferrario et al., 2016).

In (Stergiopoulos et al., 2016), authors proposed to use graphs for time-based critical infrastructure dependency analysis for large-scale and cross-sectoral failures in CI. From a supply chain perspective, many approaches study the dynamics of the interdependence between CI and suppliers to identify, assess, and mitigate risks within their end-to-end supply chain, thus improve CI resilience (Hossain et al., 2019; Trucco et al., 2018). Modeling and assessing interdependencies either between CIs or CI and suppliers can identify and minimize the cascading risk to CI but does not directly address the CI resilience & reliability.

Existing work focusing on individual CI mostly delves into dynamically assessing industry IT and ICT networks by evaluating the cascading failures over time between assets involved in and among different business processes (Stergiopoulos et al., 2017), utilizing graphical models over the system architecture and perform risk analyses to understand ICT (i.e., PLC, RTU, SCADA) weaknesses in the industry (Cherdantseva et al., 2016; Johansson et al., 2009) and performing targeted, technical attacks on individual ICS systems; e.g., binary manipulation of ladder logic in PLCs, attacking actuator software etc. (Adepu et al., 2020; Ylmaz et al., 2018). Other approaches utilize the concept of security-by-design to provide more flexible and effective ways to secure ICT/OS solutions during software development (Cavoukian & Dixon, 2013; Filkins, 2020). 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 involving in the production chain of a CI.

From an engineering perspective, the concept of multicriteria optimization in infrastructure design is not new; however, most of them focus on the topic of cost-effectiveness (Chen et al., 2018; Rizki et al., 2020). Also, 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 impact) (Funke & Becker, 2020; Page & Wohlgemuth, 2010). These studies in the general area of infrastructure design optimization do not consider the security perspective of CI. Other techniques implement security-by-design during the implementation stage by selecting certified components that are inherently secure-based on specific cybersecurity standards (ANSI/ISA-62443, 2020). To that end, security-by-design should go beyond protecting the individual system components and how they are secure based on their design.

The benefits of modeling in the design process have allowed a what-if analysis of component failures and hazardous conditions in systems that are not yet implemented, thereby saving time, reducing costs while managing risk (Bakirtzis et al., 2020). Although the theory and visualization techniques for transitioning to model-based cybersecurity analysis are advancing (Dwivedi, 2018; Jauhar et al., 2015), there are still several challenges. Most approaches focus only on cyber-attacks utilizing attack vectors to assess the risk ignoring other types of threats. Also, they require system designers and engineers to be aware of possible cybersecurity threats without necessarily being security analysts themselves.

Compared with the mentioned studies available in the scientific literature, the new methodological approach proposed in this paper mainly focuses on individual critical industrial infrastructures (like energy corridors for oil and gas supply, water, and waste treatment plants). We utilize MFA and MFN (Funke & Becker, 2020; Page & Wohlgemuth, 2010) to model the underlying system. Also, we utilize a similar methodology with (Kotzanikolaou et al., 2013; Stergiopoulos et al., 2016, 2017) to model dependencies between the different components of the modeled system and assess the cascade risk due to possible disruptions considering an all-hazardous approach. Furthermore, we model the suppliers of a critical industrial infrastructures considering the dependencies to external sources (i.e., CIs, other industries) and assess the cascading impact due to such CI dependencies similar to (Kotzanikolaou et al., 2013; Min et al., 2007; Rinaldi et al., 2001; Stergiopoulos et al., 2016).

3 BUILDING BLOCKS

This engineering design methodology uses four building blocks:

1) A method that models material, energy, and informational flows of production chains in critical industrial infrastructures into a material flow network utilizing MFA principles.
2) A modeling algorithm that maps a material flow network to a risk dependency graph to map flow network components and their interdependencies based on the industrial production chain.
3) A risk calculation methodology in which we estimated the likelihood of a potential threat and its impact(s) disrupting a CI system components.
4) A multi-risk dependency analysis methodology for assessing the risk of the graph's dependency paths and the graph's overall risk.

Each building block is briefly presented below.

3.1 Material Flow Network Modelling

Our modeling approach is based on the principles of MFA (Allesch & Brunner, 2015; Arampatzis et al., 2016) and MFN (Funke & Becker, 2020; Page & Wohlgemuth, 2010), which model processes and material and energy flows in production chains. As such, we model flow networks as graphs with four different nodes: (i) processes, (ii) junctions, (iii) input, and (iv) output nodes. These nodes are connected with links (Figure 1). *Processes* represent single activities in which resources (material, energy, and information) are processed and transformed into other resources. *Input Nodes* are the initial sources of resources flowing towards processes and essentially represent different external resource suppliers (e.g., industries, CI). In contrast, *Output Nodes* are the destination of resources flowing from processes and represent different external resource receivers (e.g., environment) or consumers (e.g., industries, households, CI). *Junctions* represent storage nodes for resources within the network, connecting processes and acting as output nodes for a process and input nodes for another process. Processes and junctions for all intents and purposes represent the assets of the modelled infrastructure. *Links* represent a way/transport mode by which resources can flow between nodes (e.g., pipes, cables, roads, ships) (Figure 1). Such flows between nodes describe the rate at which resources are consumed (input flows) and produced (output flows). For modelling purposes, we characterized input flows as regular or backup, determining the consumption lifecycle. Also, input or output flows are assigned to the same link if they share the same transport modes, thus having the same failure probability due to transportation disruption.

A crucial step in the modeling procedure is the specification of a process. This involves the definition of the input and output resources and the relations between input and output flows. Processes are the principal entities of flow-networks. They describe activities that require a set of resources (input) and, generate new or modified resources as output. In our approach, as an exception to the previous rule, 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.

By utilizing the proposed methodology, we can create a model representing an actual production system of a CI. If all processes are specified accurately and, more importantly, the actual dependencies between them are also defined, then and only then we can have a holistic view of the system and understanding of the interactions and dependencies between the physical elements of the production chain to be able to assess the cascade risk due to a disruption to a flow network node regardless of the type of threat. By analyzing the cascading risk of each flow network node, we can evaluate the overall CI risk.

3.2 Risk Dependency Graph

To analyze and assess the flow network nodes' risk and evaluate the infrastructures overall risk, we need to map all material flow network nodes (input, output, junction, process nodes) as possible failure nodes, along with all the flows and dependencies between them, into a risk dependency graph. In the case of multiple flows between a pair of nodes, a single dependency is mapped. The created graph models the flow of resources as input and output dependencies from one possible failure node to another.

Dependencies are modelled in directed, weighted graphs $G = (V, E)$, where the nodes V represent the possible failure nodes of the flow network 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 3.3.

We focus on external/internal risks in input resource supply, along with asset availability risks.

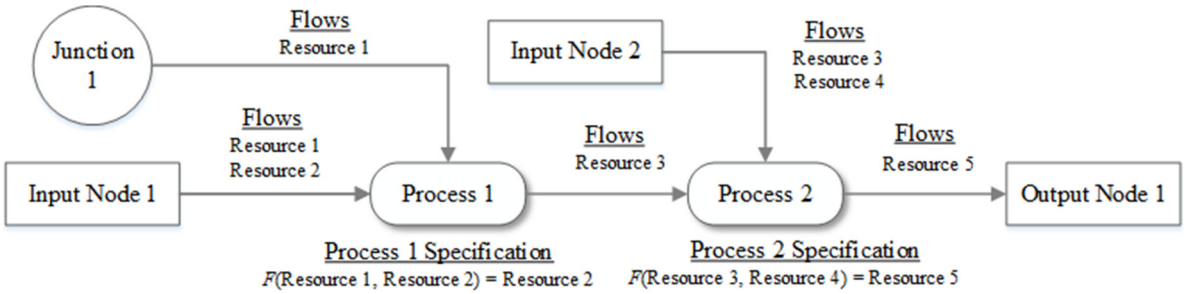

Figure 1: Graphical representation of an example material flow network.

Such risks 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., 2012), we do not differentiate between disruption types but rather consider disruptions in general and their affect in the production line. Each node in the flow network is either entirely disrupted or fully operational. This binary approach is a typical way to model disruption of resources supply (Snyder et al., 2016), while it can be used to simulate disruptions in the field of CIP (Eid & Rosato, 2016).

3.3 Risk Calculation

We calculate the risk of disruption for each possible failure node in the risk dependency graph. The standard reference of risk as a cybersecurity assessment metric is the following Eq. 1 and Eq. 2:

$$\text{Risk} = \text{Likelihood} * \text{Impact} \quad (1)$$

$$\text{Likelihood} = \text{Threat} * \text{Vulnerability} \quad (2)$$

To calculate the risk, we must first calculate the likelihood and impact values for each possible failure node in the modeled risk dependency graph. The likelihood calculation and impact evaluation are thoroughly discussed in the following subchapters.

3.3.1 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 depicts the probability that a threat can disrupt a flow network node (i.e., process, junction, input/output node), either by obstructing its activity or disrupting the supply of one or more required resources. A flow network node N can have two states: either disrupted (N) or functional ($\bar{N}$); similarly, an input resource R is either absent (R) or available ($\bar{R}$). For a single node N with required resources R, the probability of state n when resources are in the state s is $P(n|s)$, and it must hold that $\sum_n P(n|s) \leq 1$ for every instantiation/state of u. The probability $P(n|s)$ is called risk parameter, and with binary state, there are 2^n parameters to be defined for a flow network node with n input resources.

To evaluate the relationship of these disruptions between network flow nodes and input resources, we utilize a noisy-OR model (Pearl, 1988) similar to (Käki et al., 2015; K. Zhou et al., 2016). The Noisy-OR model assumes independence of causal influences among a flow network node and its required input resources (Pearl, 1988). 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 modeling challenges that MFA models introduce.

For network flow graph G with N nodes and E edges/transfer links, we define the probability $b_{i|j}$ that reflects the chance of disruption to cascade between a node i and a node j, the probability $a_i, \forall i \in N$ that reflects the chance of disruption in each node individually due to malicious or accidental activity, and the likelihood $L_i, \forall i \in N$ that reflects the chance of disruption due to the lack or not of a required resource from one or more internal or external supplier node. Therefore, to calculate the likelihood L_i of a flow network node we utilize the following formula Eq. 3 based on the noisy-OR model.

$$L_i = \sum_{\forall u} P(x|u) \quad (3)$$

For nodes without input resources (e.g. input nodes), we have $L_i = a_i$. The availability of an input resource depends on the supplier flow network nodes. To that end, we define the likelihood LR_i that reflects the availability probability of an input resource i. In case the supplier flow network nodes share the resource demand through regular flows we utilize Eq. 4 to calculate the likelihood of a resource.

$$LR_i = \sum_{\forall v} P(y|v) \quad (4)$$

For example, the calculation of the likelihood LR_i for an input resource R of a process node P that it is supplied from two flow network nodes $\{N1, N2\}$ with regular flows is shown in Table 1.

If resource a has several backup flows B_a and a likelihood value LR_a due to regular flows the final likelihood value FLR_a for that resource is calculated in Eq. 5:

$$FLR_i = \frac{LR_i}{B_i + 1} \quad (5)$$

Table 1: Resource with regular flows likelihood calculation example.

States $\boldsymbol{u}$	$\boldsymbol{P(R|u)}$
$u_1 = \{\overline{N1}, \overline{N2}\}$	$(1 - a_{N1})(1 - a_{N2})b_{P\|N1}b_{P\|N2}$
$u_2 = \{\overline{N1}, N2\}$	$(1 - a_{N1})a_{P2}b_{P\|N1}$
$u_3 = \{N1, \overline{N2}\}$	$a_{N1}(1 - a_{N2})b_{P\|N2}$
$u_4 = \{N1, N2\}$	$a_{N1}a_{N2}$

3.3.2 Impact Evaluation

Each modeled flow network node is assigned an impact value. This metric depicts the magnitude of harm due to the loss of availability, integrity of a process, or storage (junction) - including the delivery/supply of required resources. The loss of a CI asset due to an accidental or intentional incident affects all dependent CI assets in the production chain, thus the system availability and integrity. In many cases, a compromised CI asset could result in significant loss of life, casualties, material harm, environmental damage, and public service disruption.

We developed a rating scale to evaluate the severity of each impact and assign a value to each node (Table 2). Since modeled flow network nodes represent critical industrial infrastructure assets, the qualitative criteria used for the development of this scale are based on the "European Scale of Industrial Accidents" (EU/JRC, 2013) and the "Seveso Directive" (Cherrier et al., 2018). Concerning material and environmental damage, the level of impact is calculated using a logarithmic scale, and human consequences, production loss, and public disruption were approximated using the "European Scale of Industrial Accidents" (EU/JRC, 2013).

As a general guide for system disruption impact evaluation, experts must consider the number of dependent CI assets and the importance of the produced resources. By utilizing an impact scale, we enable experts to evaluate CI assets' impact based on their experience and knowledge.

3.4 Dependency Risk Analysis

Potential disruption to CI asset 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 CI assets/ components in the production chain.

To calculate and assess the n[th]-order cascading risks propagated in a series of components, we use the following method that utilizes a recursive algorithm based on (Kotzanikolaou et al., 2013; Stergiopoulos et al., 2016). Given $A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_n$ is an nth-order dependency between n connected components, with weights $R_{i,i+1} = L_{i,i+1} I_{i,i+1}$ corresponding to each 1[st] -order dependency of the path, then *the cascading risk exhibited by A_n for this component dependency path* is computed using Eq. 6:

$$R_{1,\dots,n} = (\prod_{i=1}^{n-1} L_{i,i+1}) I_{n-1,n} \quad (6)$$

The cumulative dependency risk is the overall risk exhibited by all the components in the sub-chains of the nth-order dependency. If $A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_n$ is a chain of CI asset dependencies of length n then the cumulative dependency risk, denoted as $CR_{1,\dots,n}$, is defined as the overall risk produced by an n[th]-order dependency Eq. 7.

$$CR_{1,\dots,n} = \sum_{i=2}^{n} (\prod_{j=1}^{i-1} L_{j,j+1})\, I_{i-1,i} \quad (7)$$

Eq. 7 assess the overall dependency risk as the sum of the dependency risks of the affected nodes in the chain due to a disruption realized in the source node of the dependency chain. Finally, using the total number n of all asset sub-chains (possible CI asset dependency paths) and their cumulative dependency risks, the methodology calculates the graph's overall risk G_r as the *sum of the cumulative dependency risk for each nth-order dependency* in the graph Eq. 8.

$$G_{\mathrm{r}} = \sum_{i=1}^{n} CR_{1,\dots,n} \quad (8)$$

4 ALGORITHM

The presented approach utilizes numerous methods to achieve its goals. Each step of the presented methodology utilizes a set of mapping procedures

Table 2: Impact rating scale developed based on the European Scale of Industrial Accidents.

Impact Value	System Activity Disruption	Deaths	Injuries	Material Damage (€)	Environmental Damage (€)	Public Service Disruption
5	Significant impact on overall functionality	>100	>1000	>1,000,000	1,000,000	>1 month
4	Some impact in key functions	11-100	101-1000	100,001-1,000,000	100,001-1,000,000	1 week to 1 month
3	Minor impact on overall functionality	0-10	11-100	10,001-100,000	10,001-100,000	1 day to 1 week
2	Minor impact on secondary functions	0	1-10	1-10,000	1-10,000	>1 day
1	No change in functionality	0	0	0	0	0

and algorithms. Each one provides insight into the CI production chain under analysis and outputs information to be used as input by the following step. This process uses three fundamental steps:

Step 1: Material Flow Modeling. We identify and input a critical industrial infrastructure's assets, suppliers and resource receivers, and the external or internal resource/material flows.

The model is constructed into a material flow network (graph) that exhibits the CI production chain.

Step 2: Dependency Modeling. We map the previously produced material flow network into a risk dependency graph. Also, we assess the risk of disruption for each CI asset based on the initial material flow network.

Step 3: Dependency Risk Analysis. The algorithm pre-computes all n-order dependencies using the asset dependency graph. Then for each dependency chain, outputs the cumulative dependency risk of each disruption path. Finally, the algorithm calculates the overall risk of the organization's connected assets (i.e., the entire network/graph risk).

5 EXPERIMENTS

To demonstrate the applicability and validate this approach, we developed a tool and modeled a part of the production chain corresponding to Liquefied Petroleum Gas (LPG) purification of an actual oil refining plant. We assess the flow network under a high-risk scenario, thereby establishing a baseline model. Next, we employ selected risk mitigation actions by altering and modifying the baseline network flow model, thus creating a redesigned model. We assess the new model and compare the results. Our aim here is not to evaluate the high-risk scenario or the type of the applied risk mitigation measures but to evaluate whether, and to what degree, our approach can indicate a risk reduction solution. Finally, we compare and discuss the results.

5.1 Tool Implementation

The framework was developed as a distributed application, 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 modeling functionalities and the preliminary risk analysis. The web application back-end is developed in Java Spring using the Neo4j graph ("Neo4j Graph Database," 2000) and handles the risk dependency analysis. The desktop application front-end is communicating and interacting with the web application back-end through an application programming interface (API).

5.2 Industrial CI Testbed

The critical infrastructure understudy corresponds to a typical Liquefied Petroleum Gas (LPG) purification unit encountered in all oil-refinery industries (Bahadori, 2014). Liquefied 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 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 (simple distillation column that separates components based on their boiling point), and (iii) the removal of sulfur compounds like hydrogen sulfur (H2S) and mercaptan (CH4S) in Amine Absorber Units (AAU). The material flow network model used in this study has been motivated by a real oil-refinery located in Western Asia, restructures to be representative of any typical LPG purification unit. For security considerations, the company name and all related data and component names were anonymized and sanitized.

5.3 Baseline Model

We identify 21 internal processes, 3 internal junctions, 4 internal inputs, and 1 external input for the part of the production line under study. Tables 3 and 4 display the flow network nodes and their respective flows. Flow network nodes depicted use generic terms and IDs in the examples below. The tool automatically maps the material flow network into a risk dependency graph (Figure 2). Each material flow network node and its respective input and output flows is used to model the asset dependency graph.

To assess the flow network model under a high-risk scenario, we choose a baseline failure rate (independent disruption probability) of 40% ($a_i = 0.4$) common to all components, and a possibility of a disruption to cascade between a node i to a node j is 80% ($b_{i|j} = 0.8$), common to all links. These values are chosen at the beginning of the modeling process and is fixed on a median taken from historical data.

Based on that, we calculate the likelihood of disruption of each node in the dependency graph (see Table 3) based on the method proposed in section 3.3.1.

Also, we assigned the impact values of each node (see Table 3) based on the methodology proposed in section 3.3.2 and information provided by the company. For example, CDU Debutanizers get a high impact value because their operation considerably affects the production process, and potential destruction can cause significant damages due to the flammable materials being processed. On the other hand, the LPG transfer pump gets a low impact value because it has a minor effect on the production process, and a potential accident can cause a limited extent of damages. The calculated risk of the first order dependencies is depicted in Table 4; we should note here that in this model, each resource flow corresponds to one dependency.

Table 3: Flow network nodes with impact-likelihood values.

Flow Network Nodes	ID	Impact	Likelihood
Debutanizer A1	P1	4	0.76
Debutanizer A2	P2	4	0.76
Debutanizer A3	P3	4	0.76
Debutanizer B1	P4	4	0.76
Amin Supply	I1	4	0.40
LPG Production A	I2	4	0.40
LPG Production Unit B	I3	4	0.40
LPG Production Unit C	I4	4	0.40
Energy Grid	I5	4	0.40
Debutanizer B2	P5	4	0.76
Debutanizer B3	P6	4	0.76
Debutanizer B4	P7	4	0.76
Deethanizer B3	P8	3	0.82
Deethanizer B4	P9	3	0.82
AAU 1	P10	3	0.84
AAU 2	P11	3	0.88
AAU 3	P12	3	0.88
AAU 4	P13	3	0.88
AAU 5	P14	3	0.88
AAU 6	P15	3	0.88
LPG Output	O1	4	0.32
LPG Transfer Pump	P16	1	0.76
Monitor & Control	P17	5	0.40
Debutanizer B5	P18	4	0.76
Debutanizer C1	P19	5	0.76
Debutanizer C2	P20	5	0.76
Tank 1	S1	5	0.40
Tank 2	S2	5	0.40
Tank 3	S3	5	0.40
Transfer Pump 1	P21	1	0.76

Table 4: Resources flows with 1-order dependency risks.

Source	Destination	Resource	Risk
I1	P10, P11, P12, P13, P14, P15	DEA	1.6
I2	P1, P2, P3	LPG+C5+H2S	1.6
I3	P4, P5, P6, P7, P18	LPG+C5+H2S	1.6
I4	P19, P20	LPG+C5+H2S	1.6
I5	P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11,P12, P13, P14, P15, P16, P17, 18, P19, P20, P21	Electricity	1.6
P1	S1	LPG+H2S	3.05
P10	S3	LPG	2.52
P11	S3	LPG	2.63
P12	S3	LPG	2.63
P13	S3	LPG	2.63
P14	S3	LPG	2.63
P15	S3	LPG	2.63
P16	O1	LPG	0.76
P17	P13, P3, P1, P18, P6, P20, P2, P12, P5, P4, P16, P10, S2, P21, P8, P15, S3, P14, S1, P11, P7, P19, P9	Monitor & Control Data	2
P18	P15	LPG+H2S	3.05
P19	S2	LPG	3.81
P2	S1	LPG+H2S	3.05
P20	S2	LPG	3.81
P21	S3	LPG	0.76
P3	S1	LPG+H2S	3.05
P4	P14	LPG+H2S	3.05
P5	P13	LPG+H2S	3.05
P6	P8	LPG+H2S	3.05
P7	P9	LPG+H2S	3.05
P8	P11	LPG+H2S	2.46
P9	P12	LPG+H2S	2.46
S1	P10	LPG+H2S	2
S2	P21	LPG	2
S3	P16	LPG	2

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 the top 10 highest risk dependency paths according to each one's total cumulative risk.

Thirty network flow nodes produced more than 444 dependency chains with orders ranging from two to six and with potential risk values between 0.76 and 7.84. System engineers and security experts can use this step's output to identify system components with potential risk above a specified threshold value. The threshold parameter is subjective; a decision-maker

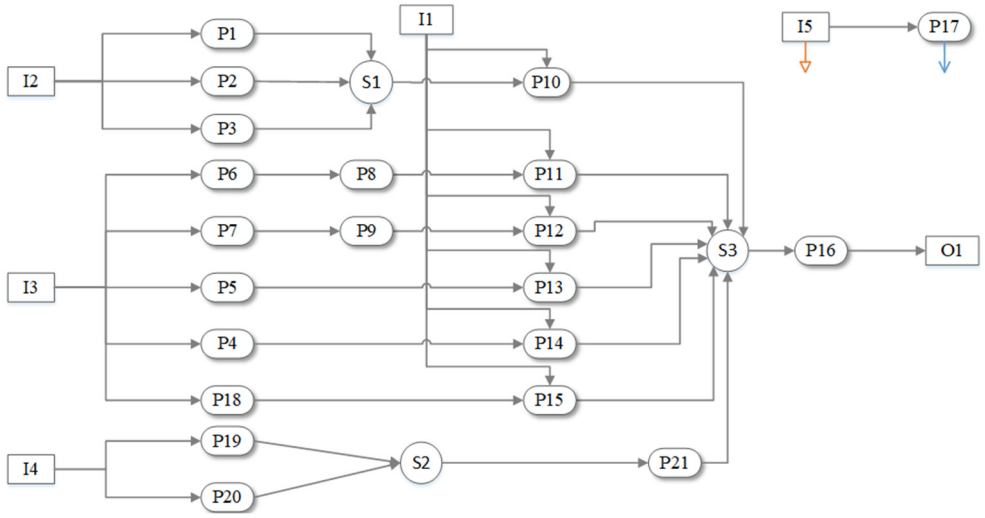

Figure 2: Tool graphical representation of the produced dependency graph (dependencies for nodes I5 and P17 have been excluded for clarity reasons).

can define it based on the critical industrial infrastructure-under-assessment or under-design specific characteristics.

Table 5: Top 10 flow network node dependency paths output from the risk analysis step (ascending).

Dependency Paths	Overall Path Risk
P7 → P9 → P12 → S3 → P16 → O1	7.84
P6 → P8 → P11→ S3 → P16 → O1	7.84
P7 → P9 → P12 → S3 → P16	7.67
P6 → P8 → P11 → S3 → P16	7.67
P18 → P15 → S3 → P16 → O1	6.6
P4 → P14 → S3 → P16 → O1	6.6
P5 → P13 → S3 → P16 → O1	6.6
P7 → P9 → P12 → S3	6.58
P6 → P8 → P11 → S3	6.58
P18 → P15 → S3 → P16	6.4

5.4 Redesigned Model

Our primary goal here is to reduce the risk of the worst dependency path and improve the overall graph risk. To achieve this, *we can either add or remove flows or network nodes to directly affect risk*; a common design decision in flow networks engineering.

However, it is not easy to justify changes. If we remove nodes or flows to reduce risk, we must be sure that factory production is unaffected. For example, if we took away all nodes, we would not have risk, but the critical infrastructure wouldn't fulfill its purpose.

For this proof-of-concept, we decided to present the safest design process available, i.e. to add backup flows και nodes for risk mitigation. To that end, we introduce one process node and five additional flows to the baseline flow network model. Tables 6 and 7 display the added flow network node and flows. The added process node corresponds to an electricity generator that acts as a backup solution for the operation and control process. All the added flows are marked as a backup decreasing the probability of disruption for the resource/material they are backing up without creating an immediate dependency.

Table 6: Backup flow network nodes and node ID association with its respective impact and likelihood values.

CI Asset	ID	Impact	Likelihood
Generator	P22	1	0.4

Table 7: Backup flows and their respective resources.

Source	Destination	Resource
I2	P6	LPG+C5+H2S
I4	P7	LPG+C5+H2S
P7	P8	LPG+H2S
P6	P9	LPG+H2S
P22	P17	Electricity

The tool produces the risk dependency graph based on the redesigned flow network and calculates each node's risk using the same probabilities of independent disruptions and disruptions to cascade under the same risk scenario as the baseline model. Finally, utilizing the risk dependency graph, the tool computes the complete set of risk paths on the risk dependency graph (Table 8). Thirty-one asset nodes produced more than 444 dependency chains with orders ranging from two to six and with potential risk values between 0.76 and 6.88.

Table 8: Top 10 CI flow network nodes dependency paths output from the risk analysis step (ascending) after the applied mitigation controls.

Dependency Paths	Overall Path Risk
P7 → P9 → P12 → S3 → P16 → O1	6.88
P6 → P8 → P11 → S3 → P16 → O1	6.88
P7 → P9 → P12 → S3 → P16	6.74
P6 → P8 → P11 → S3 → P16	6.74
P18 → P15 → S3 → P16 → O1	6.6
P4 → P14 → S3 → P16 → O1	6.6
P5 → P13 → S3 → P16 → O1	6.6
P5 → P13 → S3 → P16	6.4
P18 → P15 → S3 → P16	6.4
P4 → P14 → S3 → P16	6.4

5.5 Discussion

The redesigned model produced an overall graph risk of 1298 with an average risk per dependency path of 2.92, while the baseline model produced an overall graph risk of 1404 with an average dependency path risk of 3.16. Based on the above results, the redesigned model performs 7.5% better than the baseline in terms of risk. The applied risk mitigation measures reduced the risk of the worst dependency path by over 12%. Also, introducing an additional process with a backup flow in the redesigned model did not increase the number of produced dependency paths. From our experiments, if the risk mitigation measures include the addition of a process with regular flows (i.e., a flow that transfers resources continuously from parent to target node), the number of the dependency paths increases exponentially. There is a trade-off between adding processes or junctions to reduce the overall graph risk and introducing unnecessary complexity by increasing the dependency paths.

From a cybersecurity perspective, the operation and control processes/systems are crucial, as they are more open and more vulnerable to cyber-attacks due to existing vulnerabilities (Johansson et al., 2009; Miller & Rowe, 2012). We observed that the monitor and control process is not part of any high-risk dependency path. The reason for the low-risk values for the operation and control process is that all the flow network entities (i.e., process, junction, input, output node) share the same independent probability of disruption. A future consideration to overcome this is to define the disruption probabilities individually based on each flow network node specific characteristics.

Overall, the methodology successfully identified the reduction in the overall risk of the redesigned model, proving that the limited risk mitigation measures that we applied reduced the overall graph risk and the risk of the worst dependency path.

6 CONCLUSIONS

In this work, we propose a risk-based dependency method focusing on the individual system components to analyse disruptions in critical industrial infrastructures. The proposed approach can model a critical industrial infrastructure's underlying assets and the interactions between them as a material flow network providing a holistic view of the system and a better understanding of dependencies between the production chain's physical elements. The methodology and the 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. Preliminary tests in a part of the production line of an existing critical industrial infrastructure show that the presented approach is effective and trustworthy.

Our approach supports the proactive study of critical industrial infrastructures with large-scale production chain dependency scenarios advancing the concept of security-by-design in critical infrastructure protection. In particular, the tool helps engineers, security experts, and decision-makers to assess dependency risks before a threat is realized. Users can identify potential hotspots and project their cascading effects by analysing the complete set of potential dependency paths. By identifying potential hotspots, they can apply countermeasures early in the project lifecycle, during the design stage, improving system reliability and resilience.

Also, a user can identify and target specific nodes to make them more reliable or improve their resilience. In this way, it is possible to evaluate various alternate mitigation measures and provide convincing arguments about the expected benefits. Users can also model and compare different designs for particular system specifications, thus creating alternate scenarios. This allows for a what-if analysis based on security criteria besides cost-effectiveness and demand management criteria that system designers and engineers ordinarily use.

6.1 Restrictions and Future Work

The presented approach has certain limitations. It applies only to critical industrial infrastructures, thus not covering the whole spectrum of CI activities. Similar to other empirical risk approaches that analyse dependencies, it relies on previous risk assessments and expert knowledge on related industries and physical components to evaluate impact. Also, while this approach can identify paths and flow network nodes as

high-risk items, it is challenging to decide the right mitigation measures to reduce those risks. Choosing to add flows and flow network nodes to address risk increases the number of dependency paths and the overall flow network risk; on the other hand, removing flows and flow network nodes certainly reduces risk but affects the system's production capacity.

Future work will focus on analysing dependency paths and specifying dependencies for applying mitigation controls to reduce overall network risk. It would also be beneficial for engineers and security officers to provide indicative solutions for risk mitigation to reduce the overall system risk. Also, future work should concentrate on overcoming the requirement for previous risk assessments.

ACKNOWLEDGMENTS

This work has been partially supported by a grant ("*Research in the Cybersecurity Domain: National Cybersecurity Strategy 2020-25*") awarded to the *Research Centre* of the *Athens University of Economics & Business* (RC/AUEB). Authors express their sincere appreciation to the *Ministry of Digital Governance* of Greece.

REFERENCES

Adenso-Diaz, B., Mena, C., García-Carbajal, S., & Liechty, M. (2012). The impact of supply network characteristics on reliability. *Supply Chain Management: An International Journal*, *17*(3), 263–276.

Adepu, S., Palleti, V. R., Mishra, G., & Mathur, A. (2020). Investigation of Cyber Attacks on a Water Distribution System. In J. Zhou, et al. (Eds.), *Applied Cryptography and Network Security Workshops,* 274–291, Springer.

Allesch, A., & Brunner, P. H. (2015). Material Flow Analysis as a Decision Support Tool for Waste Management: A Literature Review: MFA for Waste Management. *Journal of Industrial Ecology*, *19*(5), 753–764.

Ani, U., Mc Watson, J., Nurse, J., Cook, A., & Maples, C. (2019). A review of critical infrastructure protection approaches: Improving security through responsiveness to the dynamic modelling landscape. *Living in the Internet of Things*, 6 (15 pp.)-6 (15 pp.).

ANSI/ISA-62443. (2020). *ANSI/ISA-62443-3-2-2020, Security for industrial automation and control systems, Part 3-2: Security risk assessment for system design*. International Society of Automation.

Arampatzis, G., Angelis-Dimakis, A., Blind, M., & Asimacopoulos, D. (2016). A web-based Toolbox to support the systemic eco-efficiency assessment in water use systems. *Journal of Cleaner Production*, *138*, 181–194.

Axelsson, J., & Kobetski, A. (2018). Towards a risk analysis method for systems-of-systems based on systems thinking. *2018 Annual IEEE International Systems Conference*, 1–8.

Azzini, I., Dido, M., Giannopoulos, G., & Galbusera, L. (2018). *GRRASP: Version 3.1: User manual.* Publications Office.

Bahadori A. (2014). Liquefied Petroleum Gas (LPG) Recovery. Bahadoni A. (Eds.), *Natural Gas Processing*, 547-590, Gulf Professional Publishing.

Bakirtzis, G., Ward, G., Deloglos, C., Elks, C., Horowitz, B., & Fleming, C. (2020). Fundamental Challenges of Cyber-Physical Systems Security Modeling. *2020 50th Annual IEEE-IFIP International Conference on Dependable Systems and Networks*, 33–36.

BS ISO/IEC 27001 (2013). *Information Technology-Security Techniques-Information Security Management Systems-Requirements.* BSI.

Cavoukian, A., & Dixon, M. (2013). *Privacy and security by design: An enterprise architecture approach*. Information and Privacy Commissioner of Ontario, Canada.

Chen, Y., He, L., Li, J., & Zhang, S. (2018). Multi-criteria design of shale-gas-water supply chains and production systems towards optimal life cycle economics and greenhouse gas emissions under uncertainty. *Computers & Chemical Engineering*, *109*, 216–235.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., & Stoddart, K. (2016). A review of cyber security risk assessment methods for SCADA systems. *Computers & Security*, *56*, 1–27.

Cherrier, V., Corden, C., Calero, J., Mazri, C., & Quintero, F. (2018). *Review of the monitoring system under the Seveso-III Directive, including the development of indicators* [Interim Report]. Wood.

Dwivedi, A. (2018). Implementing Cyber Resilient Designs through Graph Analytics Assisted Model Based Systems Engineering. *2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 607–616.

Eckert, C., & Isaksson, O. (2017). Safety Margins and Design Margins: A Differentiation between Interconnected Concepts. *Procedia CIRP*, *60*, 267–272.

Eid, M., & Rosato, V. (2016). Critical Infrastructure Disruption Scenarios Analyses via Simulation. In R. Setola, et al. (Eds.), *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach* (pp. 43–61). Springer International Publishing.

EU/JRC. (2013). *Corrosion-related accidents in petroleum refineries: Lessons learned from accidents in EU and OECD countries.* Publications Office.

Evaluation study of Council Directive 2008/114. (2020). *On the identification and designation of European critical infrastructures and the assessment of the need to improve their protection*. EU Publications.

Ferrario, E., Pedroni, N., & Zio, E. (2016). Evaluation of the robustness of critical infrastructures by hierarchical graph representation, clustering and Monte Carlo simulation. *Reliability Eng. & System Safety*, *155*, 78–96.

Filkins, B. (2020). *Security by Design: A Systems Road Map Approach.* SANS.

Funke, T., & Becker, T. (2020). Complex networks of material flow in manufacturing and logistics: Modeling, analysis, and prediction using stochastic block models. *Journal of Manufacturing Systems*, *56*, 296–311.

Hossain, N., Jaradat, R., Marufuzzaman, M., Buchanan, R., & Rianudo, C. (2019). Assessing and enhancing oil and gas supply chain resilience: A Bayesian network-based approach. *IIE Annual Conference*, 241–246.

Hulse, D., Hoyle, C., Goebel, K., & Tumer, I. (2019). Quantifying the Resilience-Informed Scenario Cost Sum: A Value-Driven Design Approach for Functional Hazard Assessment. *Journal of Mechanical Design*, *141*(2).

ISACA. 2012b. (2012). *COBIT 5: A business framework for the governance and management of enterprise IT*. ISACA.

Jauhar, S., Chen, B., Temple, W., Dong, X., Kalbarczyk, Z., Sanders, W., & Nicol, D. (2015). Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios. *2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing*, 319–324.

Johansson, E., Sommestad, T., & Ekstedt, M. (2009). Issues of cyber security in SCADA-systems-On the importance of awareness. *IET Conference Publications*, 969–969.

Käki, A., Salo, A., & Talluri, S. (2015). Disruptions in Supply Networks: A Probabilistic Risk Assessment Approach. *Journal of Business Logistics*, *36*(3), 273–287.

Kobetski, A., & Axelsson, J. (2017). Towards safe and secure systems of systems: Challenges and opportunities. *Proc. of the Symposium on Applied Computing*, 1803–1806.

Kotzanikolaou, P., Theoharidou, M., & Gritzalis, D. (2013). Assessing n-order dependencies between critical infrastructures. *International Journal of Critical Infrastructures*, *9*(1/2), 93.

Miller, B., & Rowe, D. (2012). A survey SCADA of critical infrastructure incidents. *Proc. of the 1st Annual Conference on Research in Information Technology*, 51.

Min, H.-S., Beyeler, W., Brown, T., Son, Y., & Jones, A. (2007). Toward modeling and simulation of critical national infrastructure interdependencies. *IIE Transactions*, *39*(1), 57–71.

Neo4j Graph Database. (2000). *Neo4j Graph Database Platform*. https://neo4j.com/product/neo4j-graph-database/

NIST SP 800-30. (2012). *Guide for conducting risk assessments* (NIST SP 800-30r1). National Institute of Standards and Technology.

Page, B., & Wohlgemuth, V. (2010). Advances in Environmental Informatics: Integration of Discrete Event Simulation Methodology with ecological Material Flow Analysis for Modelling eco-efficient Systems. *Procedia Environmental Sciences*, *2*, 696–705.

Parfomak, P. (2007). *Vulnerability of concentrated critical infrastructure: Background and policy options*.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Elsevier.

Rinaldi, S., Peerenboom, J., & Kelly, T. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems*, *21*(6), 11–25.

Rizki, Z., Janssen, A., Claassen, G., Boom, R, van der Padt, A. (2020). Multi-criteria design of membrane cascades: Selection of configurations and process parameters. *Separation and Purification Technology*, *237*, 116349.

Setola, R., Luiijf, E., & Theocharidou, M. (2016). Critical Infrastructures, Protection and Resilience. In R. Setola, et al. (Eds.), *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach,* 1–18, Springer.

Snyder, L., Atan, Z., Peng, P., Rong, Y., Schmitt, A., & Sinsoysal, B. (2016). OR/MS models for supply chain disruptions: A review. *IIE Transactions*, *48*(2), 89–109.

Stergiopoulos, G., Kotzanikolaou, P., Theocharidou, M., Lykou, G., & Gritzalis, D. (2016). Time-based critical infrastructure dependency analysis for large-scale and cross-sectoral failures. *International Journal of Critical Infrastructure Protection*, *12*, 46–60.

Stergiopoulos, G., Kouktzoglou, V., Theocharidou, M., & Gritzalis, D. (2017). A process-based dependency risk analysis methodology for critical infrastructures. *International Journal of Critical Infrastructures*, *13*(2/3), 184.

Trucco, P., Petrenj, B., & Birkie, S. E. (2018). Assessing Supply Chain Resilience upon Critical Infrastructure Disruptions: A Multilevel Simulation Modelling Approach. In Y. Khojasteh (Ed.), *Supply Chain Risk Management: Advanced Tools, Models, and Developments*, 311–334, Springer.

Ylmaz, E. N., Ciylan, B., Gonen, S., Sindiren, E., & Karacayilmaz, G. (2018). Cyber security in industrial control systems: Analysis of DoS attacks against PLC and the insider effect. *2018 6th International Istanbul Smart Grids and Cities Congress and Fair*, 81–85.

Zhou, K., Martin, A., & Pan, Q. (2016). The Belief Noisy-OR Model Applied to Network Reliability Analysis. *International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems*, *24*(06), 937–960.