Critical Infrastructure Automated Immuno-Response System (CIAIRS)

By

Sahar Khalid Badri

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy

"If you said I can't… You will never be lying… never ever limit yourself... "

This Dissertation is Dedicated to the Memories of My Brother,

Ahmed k. Badri (1986-2014)

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

"In the Name of Allah, the Most Beneficent, the Most Merciful"

I would like to start by thanking "ALLAH"; I wouldn't be able to finish this journey without his generous help. I had a lot of ups and downs but my faith in "Allah" was bigger and I was sure he would never leave me alone. Finally, I can say "one of my dreams came true!"

I would like to express my sincere gratitude to my advisors Dr Paul, William and Nathan for their continued support of my PhD study and related research, for their patience, motivation, and immense knowledge. I would particularly like to thank William for his guidance throughout this research and the writing of this thesis. I could not have imagined having a better advisor and friend for my PhD study than you.

I would like to thank my family: my parents, brothers, sisters and my soul mate Amr, for supporting me spiritually throughout this work. To my parents– we have experienced some difficulties in the past three years. Every time I was ready to quit you did not let me and I am forever grateful. This dissertation stands as a testament to your unconditional love and encouragement. Also, I will never forget my little kids Malak and Khalid. They have started this journey with me and have supported me and handled all my tensions. Thank you, Malak and Khalid, for being always there for me and I will not forget my little new baby Nour, you were "icing on top of the cake"!

Nevertheless, at the end of this journey, I have gained friends that I will never forget. Tricia is one of the kindest people that I have met, I will really miss you!

Also would like to thank all the technicians for their kind help especially Neil, Mick and Ian. They will never close their doors for any student!

ABSTRACT

Critical Infrastructures play a central role in the world around us and are the backbone of everyday life. Their service provision has become more widespread, to the point where it is now practically ubiquitous in many societies. Critical Infrastructure assets contribute to the economy and society as a whole. Their impact on the security, economy and health sector are extremely vital. Critical Infrastructures now possess levels of automation that require the integration of, often, mutually incompatible technologies. Their increasing complexity has led to the creation of direct and indirect interdependent connections amongst the infrastructure groupings. In addition, the data generated is vast as the intricate level of interdependency between infrastructures has grown.

Since Critical Infrastructures are the backbone of everyday life, their protection from cyber-threats is an increasingly pressing issue for governments and private industries. Any failures, caused by cyber-attacks, have the ability to spread through interconnected systems and are a challenge to detect; especially as the Internet is now heavily reliant on Critical Infrastructures. This has led to different security threats facing interconnected security systems. Understanding the complexity of Critical Infrastructure interdependencies, how to take advantage of it in order to minimize the cascading problem, enables the prediction of potential problems before they happen.

Therefore, this work firstly discusses the interdependency challenges facing Critical Infrastructures; and how it can be used to create a support network against cyber-attacks. In much, the same way as the human immune system is able to respond to intrusion. Next, the development

of a distributed support system is presented. The system employs behaviour analysis techniques to support interconnected infrastructures and distribute security advice throughout a distributed system of systems. The approach put forward is tested through a statistical analysis methodology, in order to investigate the cascading failure effect whilst taking into account the independent variables. Moreover, our proposed system is able to detect cyber-attacks and share the knowledge with interconnected partners to create an immune system network. The development of the *'Critical Infrastructure Auto-Immune Response System*' (CIAIRS) is presented with a detailed discussion on the main segments that comprise the framework and illustrates the functioning of the system. A semi-structured interview helped to demonstrate our approach by using a realistic simulation to construct data and evaluate the system output.

INDEX OF TERMS

Critical Infrastructure, Critical Infrastructure Protection, Behaviour Observation, Classification, Interdependency, Statistical Analysis Methods, Big Data, Distributed System, System of Systems, Data Analysis, Cyber-attack, Simulation, Immune System, Statistical Descriptive, Dataset, SCADA, UTM.

GLOSSARY

- Asset: A hardware or software component that needs to be protected from attacks.

- Artificial Immune Systems (AIS): Solving real-world problems by intelligent methodologies inspired by the organic immune system.

- Behavioural Observation: A technique that is used to observe, evaluate and calculate the behaviour of a system.

- Big Data: Extremely large blocks of data that can be used to discover system behaviour by analysing patterns, trends and different information related to the system.

- Control Centre: When a control system is managed by an operational centre. The control centre includes SCADA and HMI systems.

- Critical Infrastructure (CI): The arrangement of both systems and assets, which are essential and affect the security, economy, public health or safety of a nation.

- Critical Infrastructure Protection (CIP): The awareness and the action toward an attack in the Critical Infrastructure.

- Cyber-Attack: Digital attacks that are attempted by a hacker to compromise information by hacking a network or information system.

- Cyber-Security: Using measures to protect information systems or a network from cyber-attacks.

- Data Analytics: Concerned with improving the productivity of a system by qualitative and quantitative techniques in order to draw a conclusion about a set of data.

- Descriptive Statistic: Refers to data analysis by summarizing or describing the number of features; such as the mean, standard deviation and mode.

- Environmental Protection Agency (EPA): An American agency established in order to protect the environmental health.

- Human Machine Interface (HMI): The process of an industrial control system through a user interface. The transferred information is managed and monitored by a control system operator.

- Industrial Control System (ICS): Refers to different types of control systems. In addition, it relates to an industrial process that is operated by systems, devices and networks.

- Master Station (MS): Involved in the industrial protocol communication session by monitoring the asset. Moreover, the station handles time management and synchronisation.

- Modbus: The Modicon Bus protocol, responsible for communication between the industrial control assets.

- Programmable Logic Controller (PLC): Is an industrial device that is responsible for collecting data from the input and output to create an automated control loop. These loops are created with logical programs.

- Remote Terminal Unit (RTU): Is a device that uses a remote location process in order to combine remote communication programs with logical programs.

- Supervisory Control and Data Acquisition (SCADA): Monitors and controls an infrastructure's operation. The system consists of two main parts: MTU and RTU. The MTU (Master Terminal Unit) acts as the brain of the system and the RTU (Remote Terminal Units) collects the data locally and sends it to the MTU. The Internet has increased the number of external connections to SCADA systems, making them more vulnerable to cyber-attack.

- Simulation: A method for teaching students to imitate elements of real-world processes, overriding difficulties, such as material cost or human resources.

- Simulation Environment: Using software simulation to imitate real systems.

- System of Systems: This is a collection of connected tasks or interdependent systems, which form part of a larger and more complex system.

- Tecnomatix: A highly comprehensive simulation program developed by Siemens. Used to evaluate the proposed system, from the process layout and the constructed of product lifecycle management.

- Threats: Something/someone that can exploit a vulnerability (both intentionally and accidentally) to obtain, damage or destroy an asset.

- Vulnerability: A weakness in a system that can be exploited by attackers to damage or break into the system.

PUBLICATIONS RESULTING FROM THIS THESIS

1. **Badri, S**., Fergus, P., & Hurst, W., "A Cyber-Support System for Distributed Infrastructures", in *EMERGING 2016 : The Eighth International Conference on Emerging Networks and Systems Intelligence*, Venice, Italy: Iaria, pp. 1-6, 2016.

2. **Badri, S**., Fergus, P., & Hurst, W., "Critical Infrastructure Automated Immuno-Response System (CIAIRS)", in *3rd International conference on control, Decision and Information Technologies (CoDIT).* Malta: IEEE, April 2016, pp. 1-6.

3. **Badri, S**., Fergus, P., & Hurst, W., "A Support Network for Distributed Systems", in *10th International Conference on E-Learning and Games, Edutainment 2016*. Springer: Hangzhou, China, April 2016, pp.16-22.

4. **Badri, S**., Fergus, P., & Hurst, W., "Statistical Analysis Methods for Interdependency Communication in Distributed Systems", in *International Conference on Developments in eSystems Engineering*. IEEE: Liverpool, UK, September 2016, pp. 1-6.

5. **Badri, S**., Abuelmaatti, O., Llewellyn-Jones, D., & Merabti, M., "An Artificial Immune System Techniques for Critical Infrastructure Support", in *14th Annual PostGraduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting (PGNet 2013)*. Liverpool, UK, July 2013, pp. 1-6.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Foreword

Over the last decade, Critical Infrastructure complexity has raised a number of challenges and prompted the use of intelligent software and simulations to address and achieve effective cyber-security solutions [1]. However, the breadth of technologies, organisations and interoperating systems that comprise Critical Infrastructures, makes security particularly challenging [2]. Therefore, new techniques are required to investigate innovative areas of technology in order to safeguard critical systems against a growing cyber-threat [3]. Considerable effort has been expended on the protection of Critical Infrastructures; it is still an ongoing and persistent challenge [4]. One significant challenge is overcoming the lack of understanding about the interdependency scheme within Critical Infrastructure groupings. Moreover, there is no single approach about how the elements of a Critical Infrastructure affect a connected partner [5] [6].

As such, this research takes the concept of a human immune system to simplify and predict potential problems before they spread through a network of infrastructures. Therefore, this chapter starts by introducing Critical Infrastructures, highlighting the motivation behind the research. In

addition, the aims and objectives are presented, as are the novelties of the research. The chapter is concluded with an overview of the thesis structure.

1.2 Critical Infrastructures

During the last century, a massive growth in Critical Infrastructure groupings has occurred, leading to issues, such as interdependency [7]. Researchers and organisations were traditionally focused on exploring and understanding Critical Infrastructure behaviour to make the system more productive rather than understanding the different threats and attacks that can influence the system and the different modelling features that need to be studied in order to explore the Critical Infrastructures interconnectivity [8]. Therefore, this section presents a brief background of the Critical Infrastructures and the interdependency challenge.

1.2.1 Brief of Critical Infrastructures

Critical Infrastructures play a significant role in the world around us. Their service provision has become more widespread, to the point where it is ubiquitous in many societies [9]. To maintain continuous supply, infrastructure interconnectivity has become highly complex; particularly due to the increase in demand for amenities. This has led to an increased interdependence between infrastructures and their underlying physical layers. One infrastructure's provision relies heavily on another [10]. Due to this increased connectivity, now, more than ever before, Critical Infrastructures face a number of possible digital threats. As a result, Critical Infrastructure Protection (CIP) has become a significant research topic [11]. Infrastructure is the main source of development and economic construction in any country, and different types of urban development depend on the size and the provision of infrastructure elements, which help guide the development of new areas [12]. Critical Infrastructures have an important influence in an urban environment

[13]. However, many infrastructures, such as power plants, are considered outdated and are difficult to repair [14].

The National Institute of Standards and Technology define Critical Infrastructures in the Executive Order (EO) as any physical or virtual system and their assets that affect the nation's security, public economy and health service by their failure or damage occurring to them [13]. Critical Infrastructure assets are also explained by Command *et al*., and can be divided into three categories. First, the physical assets, which could be tangible or intangible. Secondly, human assets, that can represent vulnerabilities by having privileged access to important information or systems. Third, cyber assets, which include hardware, software and data, which all serve the network functionality [15].

Critical infrastructure security attracts the attention of various research fields. Yusufovna *et al*., for example, discuss energy resources, finance, food, health, government services, manufacturing, law and legislation, including transportation [16]. In addition, Baud *et al*., include additional areas, such as water, information and telecoms, chemical, industry, agriculture, postal and shipping, and defence [17].

1.2.2 Brief of Interdependency

Due to global expansion of technology and with the Internet revolution, infrastructures have become highly complex and have increased the interdependency at the physical and network layers [18]. Therefore, interdependency is considered to be one of the characteristics that can raise several concerns; in particular the analysis and modelling of interdependencies due to the complicated interactions [19]. This interdependency challenge, within the Critical Infrastructure system-of-systems, has the potential to cause a cascading effect, with unprecedented disastrous outcomes.

Therefore, understanding the interconnectivity behaviour between Critical Infrastructures and how it changes depending on the complexity, can help in reducing the effect before cascading occurs [20]. This would control the damage and limit the impact [21]. Every new interdependency reveals a fresh vulnerability in the interconnected infrastructure groups, which creates new attacks risks [22].

A number of factors contribute to make the interdependency more complex. For example, timescales, geographic scales, cascading and higher order effects, social/psychological elements, operational procedures business policies, restoration and recovery procedures, government regulatory, legal, policy regimes and finally stakeholder concerns [23]. These aspects can have a detrimental impact on the system, so services need to be secured against any type of attack and mitigation plans should be in place to counter the effects of other disasters [24]. In order to achieve this, security planning must anticipate on a large-scale.

1.3 Motivation

Although there are a considerable number of studies focused on Critical Infrastructure research; many of them tend to focus on understanding the improvement of Critical Infrastructures, modelling the infrastructure, or examining different classifiers to improve the efficiency of Critical Infrastructures. However, these studies are limited by the fact that challenges related to interdependency caused by fast expansion are ignored. In addition, little research has been done to study the communication between Critical Infrastructure and how interdependency can be used as a progressive point to improve the level of security and moderate the cascading problem.

Therefore, the main motivation behind this thesis is based on two main aspects. Firstly, Critical Infrastructure tolerance towards growing cyber threats and secondly, the interdependency effect

between the Critical Infrastructures and how modelling can help to understand the challenges that interconnectivity incurs.

1.3.1 Cyber Threats Risks

One necessity in our life that cannot be dispensed with is the Internet. The Internet has an important role in the dissemination of ideas, news and policies, cultures and economic transactions. The significant expansion of the Internet has allowed criminals to exploit systems [25] [26]. Despite the remarkable growth in research to develop systems to ensure the necessary protection, attackers on the Internet work in parallel to enhance their own approaches [27].

There is no safe way to protect against intrusion, whatever the strength of protection. Therefore, it takes constant vigilance and advanced technology to mitigate the effects of cyber-attacks [28]. Billions have been spent on defending Britain against cyber-attacks and new cyber security strategies have helped, to some extent, contain the growth of criminal organisations [29]. However, understanding all the assets an organisation needs to protect against to reduce the risk of threats is impossible. Tovey *et al*., highlighted the scale of the problem and compared the average IT spending with the average revenue loss after the occurrence of cyber-attacks on different sectors in the UK. It was clear that utilities, energy and mining sectors had the second highest level in spending on IT at 6% whilst having the highest revenue loss as a result of cyber-attacks [30].

Yet Giannopoulos, *et al*., argue that security spending should remain high, to ensure ongoing protection of Critical Infrastructures, such as, the power grid, the transport network and information and communication systems and their services they provide [31]. Since any damage within the Critical Infrastructures by any means, whether accidental, such as by natural disasters,

or intentional by terrorism, and criminal activity, might have negative consequences on the security of the EU and directly affect citizens [31].

After the terrorist attacks in 2001, many works were focusing on securing the Water Critical Infrastructure. However, the Environmental Protection Agency (EPA), who are committed to protecting drinking water and wastewater utility systems, did not implement protocols that met a specific security level until 2005. Moreover, most security enhancements were implemented for free by organisations [32]. In 2010, the EPA revealed one of their developed policies for the water system, which concerns improving the water planning process and the security level of the system [33]. However, the water system infrastructure is a unique challenge due to the need to constantly repair or replace services, due to faults caused by weather changes, population growth and cyber attacks [34].

1.3.2 Interdependency Effects

As a result in the growth of the terrorist attacks on Critical Infrastructures, it is becoming important to focus on the CIP [35]. Especially given that Critical Infrastructures are becoming more complex and interconnected [36]. With this advance in complexity, a high risk of failure is predicted from one Critical Infrastructure to another. Although considerable effort has been expended on the protection of Critical Infrastructures, it is still an ongoing and persistent challenge [37]. Various reasons make this challenge hard; one of them is the lack of a method for completely understanding the interdependency scheme within Critical Infrastructure groupings [38].

Taking the impact of the 2008 Chinese winter storm on several Critical Infrastructures as an example; the storm affected several Critical Infrastructures by ice and heavy snow causing huge economic damage. However, it also showed the level of interconnectivity between electricity

power outages, transportation and communication networks and the effect this has on service delivery. Rong *et al.,* detail a data analysis process that can be used to understand the interdependency within the Critical Infrastructures [39], [40]. Therefore, understanding the links between Critical Infrastructures that simplify the structure and predict potential problems before they happen is an essential process in protecting against cascading failure.

1.4 Aims and Objectives

The inspiration behind this thesis is to detail the development of an approach called Critical Infrastructure Automated Immuno-Response System (CIAIRS), which increases the level of security for interconnected Critical Infrastructure by predicting and communicating potential cyber-attacks before they spread through a network of infrastructures. Consequently, the aims of this thesis are presented as follows:

- To investigate the extent to which Critical Infrastructures are interconnected.

- To develop an effective response to the challenges in Critical Infrastructures as a result of their interconnectivities.

- To explore how the interdependency between Critical Infrastructures affects their individual systems.

- To design a system that understands the behaviour of system data and classifies it as either normal or abnormal, depending on data distributions. The system aims to increase the level

of security for the Critical Infrastructure and to ensure Critical Infrastructures can refine and improve the way attacks are handled.

- To understand interdependency modelling methods in order to produce an adequate simulation environment that can be used for experimentation and realistic data construction.

- To demonstrate a technique for the detection of abnormal behaviour within a Critical Infrastructure and offer an approach for sharing the information with other infrastructures.

In order to achieve the above aims the following objectives need to be accomplished, which have been designated as fundamental to the research:

- Research into related literature, concerning CIP, the interdependency modelling challenge, Critical Infrastructure simulation models, Critical Infrastructure security, data analysis and data classifiers.

- Research into the impact of Critical Infrastructure interdependency and identify the influence on the interdependencies, which cause devastating damage.

- Evaluate and examine various modelling methods used for interdependency challenges.

- Find available software that covers the requirements to implement the system prototype in order to evaluate the effectiveness of the system.

- Create a case study through different simulations in order to shape the idea of the research in real life depending on semi-structured interview.

- Implement and evaluate different case studies in order to measure the performance of the simulation and the system in detecting the abnormal behaviour effects on other Critical Infrastructures.

- Examine data classifiers and detect the best data classification process, which will help in identifying the behaviour of the Critical Infrastructure pattern.

- The publication of the results obtained during the research.

1.5 Novelties

Through the process of this research, fulfilling the aims and objectives were important in order to accomplish the research's novelty. The novelties include the production of an innovative system to support the prediction of abnormal behaviour in Critical Infrastructure systems before they become serious and affect other interconnected systems. This is especially important in the context of our increasing dependence on Critical Infrastructures, along with their increasing complexity. The novelties in this work can be summarised as follows:

- The novelty of the CIAIRS design depends on understanding the behaviour of the attacks rather than using the known attack signatures. By using the behavioural observation process attacks can be generalised and predicted before they occur.

- CIAIRS is able to detect cyber-attacks and share the knowledge with interconnected partners. Potential cyber-attacks are detected and the potential impact is predicted before they spread through a network of infrastructures. This is inspired by the human immune system approach. Big data analysis techniques are used to identify and share threats between different infrastructures. This technique is a novel application since it contains a number of components that shape the idea of using behaviour observation, classification and communication links depending on the interconnectivity between Critical Infrastructures [41], [42].

- Collecting data in real-time will allow the system to be sophisticated in processing a fast response toward any new abnormal behaviour.

- Current work focuses on shielding Critical Infrastructures. Whereas this research aims to develop a solution that both shields and shares the attack information with connected partners. The communication process for sharing attack information between the different partners is one of the novelties of the design. This process is divided into three main stages: decision, inner-process and action. In addition, the use of response and recommendation processes are also novel [43].

1.6 Thesis Structure

The thesis is divided into 7 chapters and presents the research idea using a logical scientific methodology. These chapters are as follows:

- Chapter 2: Background: this chapter provides background information on three main topics, which will shape the main idea of the research. The first sub-section is the background on the Critical Infrastructure. This section focuses on defining the Critical Infrastructure, highlighting different Critical Infrastructure types and the different challenges that are faced. The second sub-section is on interdependency in Critical Infrastructures. This sub-section indicates the main four categories for interdependency, and presents some of the interconnectivity modelling examples. The final sub-section covers cyber-attacks and the challenges Critical Infrastructures face.

- Chapter 3: Literature Review: this chapter presents Critical Infrastructures, big data and the use of the analytics in prediction. In addition, the chapter indicates a pattern of communication used in the framework that is based on the publish-subscribe design methodology. Simulation is also discussed in this chapter which explains the importance of simulation and how it can be a useful tool for the advancement of Critical Infrastructure security. The chapter highlights the simulation types and the role of using it to protect Critical Infrastructures. Finally, an in-depth discussion on machine learning is presented.

- Chapter 4: Critical Infrastructure Automated Immuno-Response System (CIAIRS): this chapter provides a detailed discussion on our methodology and proposed system. Initially, the system requirements and the system location within the ICS layers are discussed. In addition, a number of scenarios are presented. Next, the system framework is presented in detail and each component is individually explained.

- Chapter 5: Implementation: this chapter applies CIAIRS to a simulated scenario based on a real-world critical infrastructure in the Saudi Water Ministry using the Siemens Tecnomatix system. This infrastructure is used to present the flow and behaviour of data in order to study the effect it has on Critical Infrastructures. The chapter explains the simulation environment and program used (Siemens Tecnomatix Plant Simulation). The simulation is used to construct normal and the abnormal data. The components and the process flow for each is described detail. The chapter also explains the best practices for presenting a productive environment. Two Critical Infrastructures, the electricity and water, are used to present the flow material. Finally, a snapshot of the normal and abnormal water data is presented.

- Chapter 6: Evaluation: This chapter present the outcome of the semi-structured interview. This chapter also discusses the classification results for normal and abnormal behaviour using data obtained through simulation.

- Chapter 7: Conclusion and Future Work: In this final chapter, the main achievements and limitations in the research are discussed. The chapter starts by stating recommendations for real-life Critical Infrastructures depending on the implication of the research findings. The chapter and thesis ends with some discussion of future work as a continuation of this research.

CHAPTER 2

BACKGROUND

2.1 Introduction

A country's foundation depends on its Critical Infrastructures. Protecting them ensures the safeguarding of the economy as the Critical Infrastructure assets contribute to the society as a whole [44]. However, their increasing complexity has led to the creation of direct and indirect interdependent connections amongst the infrastructure groupings. In addition, the datasets generated are vast and intricate [45]. Any failures, caused by cyber-attacks have the ability to spread through a system of systems and are a challenge to detect [46].

Therefore, this section, discusses the Critical Infrastructure in depth, focusing on a number of points, such as the challenges facing Critical Infrastructure interdependency.

2.2 Critical Infrastructure

This research looks into understanding the links between Critical Infrastructures, in order to simplify the systematic use of these structures and to predict potential problems before they occur. However, the interdependence between Critical Infrastructures still presents a dilemma [47]. As with any other system, Critical Infrastructures face a number of possible attacks. As such, CIP has become the focus of many researchers' attentions for some time and has become a cause of concern

to governments and private sectors around the world. As a result, CIP has become an important research topic [48][25]. In this section, a background discussion is put forward on the definition of a Critical Infrastructure, the challenges they face and the risks posed by infrastructure interconnectivity.

2.2.1 Background

A number of different definitions can be found for the term Critical Infrastructure (CI). Therefore, we choose to present two of these definitions that were comprehensive in identifying the meaning of Critical Infrastructure.

Moteff *et al*., define Critical Infrastructures in various ways, one of them being as the arrangement of both systems and assets, which are essential and can affect the security, national economic security, national public health or safety, or any combination of those matters of a society by their failure or damage [49]. A similar definition regarding Critical Infrastructure was given by Command *et al*., [15]. On the other hand, Moteff, *et al*., also indicate that Critical Infrastructures are a structure of mutually dependent networks and specific industry systems, organizations and actions and distribution capabilities that supply vital products and services for the protection and financial security of society [49].

Infrastructure is the main source of development and economic construction process of any country. Different types of urban development depend on the size and the provision of infrastructure elements, which help guide the development of new areas [50]. Critical Infrastructures have an important influence in an urban environment. However, many infrastructures, such as power plants, are considerably outdated and difficult to repair [13].

The National Institute of Standards and Technology (NIST) defines Critical Infrastructures as any physical or virtual system and assets that would affect the nation's security, public economy and health service by failure or damage occurring to them [13].

Critical Infrastructure assets are also explained by Command *et al*., and can be divided into three groups: physical assets (which could be tangible or intangible), human assets (that can represent vulnerabilities by having privileged access to important information) or systems cyber assets (which include hardware, software and data, and which all serve the network functionality) [15]. Moreover, Critical Infrastructure security attracts the attention of various research fields. The Critical Infrastructure's common areas were devolved during time and Yusufovna *et al.,* indicate these areas as follows [16]:

- Energy resources
- Finance
- Food
- Health
- Government services
- Manufacturing
- Law and legislation
- Transportation
- People and education

In addition, Baud *et al.*, include additional areas, such as water, information and telecoms, chemical, industry, agriculture, postal, shipping and the defence industry [17].

However, the relevant assets of these Critical Infrastructures differ depending on the interest of the country, geopolitical developments or the security policy in question [49]; Table 1 presents an example of the different assets.

Table 1: Assets in Critical Infrastructures (in the US) from 1983-2003 [49]

Infrastructures	U.S Government Reports and Executive Orders							
	CBO (1983)	NCPWI (1988)	E.O. 13010 (1996)	PDD-63 (1998)	E.O.13228 (2001)	NSHS (2002)	NSPP (2003)	HSPD-7 (2003)
Transportation	x	x	x	x	x	x	x	x
Water supply/waste water treatment	x	x	x	x	x	x	x	x
Education	x							
Public health	x			x		x	x	x
Prisons	x							
Industrial capacity	x							
Waste services		x						
Telecommunications			x	x	x	x	x	x
Energy			x	x	x	x	x	x
Banking and finance			x	x		x	x	x
Emergency services			x	x		x	x	x
Government continuity			x	x		x	x	
Information systems				x	x	x	x	x
Nuclear facilities					x			
Special events					x			
Agriculture/food supply						x	x	x
Defense industrial base						x	x	x
Chemical industry						x	x	x
Postal/shipping services						x	x	x
Monuments and icons							x	x
Key industry/ tech sites							x	
Large gathering sites							x	

2.2.2 Infrastructures Characteristics: Challenges

This section begins by illustrating some of the characteristics of Critical Infrastructures to help in clarifying the features, which the infrastructure holds within its systems. The infrastructure is capable of changing its geographical state depending on system expansion [51]. This has led to Critical Infrastructures being considered as an interdependent set of infrastructures [19].

There are three elements of Critical Infrastructures, which are particularly important in order to understand the behaviour of these infrastructures from a security perspective. A detailed description of these can be found in Denis *et al*., but we also summarise them here. Firstly, SCADA is a key technology for Critical Infrastructures [11]. SCADA is a communication system that

supports a collection of data from sensors and sends them to a control centre for management purposes [52]. However, by introducing SCADA into a system, new concerns are also introduced in terms of security (This point will be explained in more detail later when we discuss the problems and challenges associated with Critical Infrastructure security). Secondly, the increasing importance of communication for Critical Infrastructure has seen the introduction of IP-based networks. Finally, an important factor is the introduction of the smart-grid, which has significant ramifications in terms of security.

Although considerable effort has been expended on the protection of Critical Infrastructures, it remains an ongoing and persistent challenge [53]. For various reasons, this challenge is hard to tackle. A particular difficulty is that there is no way to fully understand the interdependencies between the various Critical Infrastructures [54]. Moreover, there is no single understanding of how the elements of Critical Infrastructure functionally affect one another [55]. Consequently, this fundamentally impacts on our ability to protect Critical Infrastructure systems, since one of the key aims is to determine and manage these elements during an attack.

Another challenge that Critical Infrastructures face is the protection of the underlying SCADA systems [56]. As discussed previously, SCADA is a system that monitors and controls Critical Infrastructure operations, such as power, water, *etc*. SCADA includes two main elements: the Master Terminal Unit (MTU) and the Remote Terminal Units (RTUs). The MTU acts as the brain of the system and the RTUs collect data locally and send it to the MTU [57]. Coutinho, *et al*., define two approaches to having a more secure and protected SCADA system. The first approach is to spot the problems that can affect the parameters in the system by using antivirus and intrusion-detection systems. The second approach is to check the data flow and the system within periods of normal behaviour and try to detect any difference or change within the SCADA system [58].

Moreover, increasing numbers of connections to SCADA systems, such as office networks and the Internet, makes them more vulnerable to conventional cyber-attacks [59].

2.3 **Interdependency**

Critical Infrastructure types vary from one country to another [60]. However, there is one aspect in common; infrastructures have become highly interdependent. This is due to the increase in demand for the amenities provided. This has led to an increased interdependence between their underlying physical layers; in which a failure can be cascaded and affect the rest of the domain. One infrastructure's provision relies heavily on another. Due to this increased connectivity Critical Infrastructures face a number of possible digital threats [61].

The Internet is heavily relied upon by the Critical Infrastructure. This has led to different security threats facing interconnected security systems [62]. By understanding the interconnectivity behaviour between Critical Infrastructures, and how it changes depending on the complexity, can help in reducing the effect before cascading occurs. This would control the damage and limit the impact [63]. Taking into account that every new interdependency reveals a fresh vulnerability in the system of systems, which creates new cyber-attack risks [49].

Considerable effort has been expended on the protection of Critical Infrastructures; however it is an ongoing and persistent challenge. Various factors contribute to this, for example, there is a lack of understanding about the interdependency scheme within Critical Infrastructure groupings. Moreover, there is no single approach about how the elements of a Critical Infrastructure's functionality are affected by a connected partner. Rinaldi *et al*., for example, identify four groups to categorise infrastructure interconnectivities [6][5].

Moreover, infrastructure interdependency can be viewed as a flow in networks. However, we cannot treat different Critical Infrastructure systems as an independent part, as it would limit our modelling methods. Four interdependency categories were identified by Rinaldi *et al*., where the infrastructure can be fitted into any of the following four types according to those categories [6] [5]:

- Physical Interdependency: it can be said that two infrastructures are physically interdependent on each other if their output materials are linked.

- Geographic Interdependency: it can be said that the infrastructures are geographically interdependent if an environmental factor can affect the infrastructure.

- Logical Interdependency: it can be said that two infrastructures are logically interdependent on each other if their connection is through a specific mechanism, such as policies, regulation, *etc*.

- Cyber Interdependency: it can be said that the infrastructures are cyber interdependent if the infrastructures depend on information transmitted through information systems. SCADA is one example of a communication system that could cause cyber interdependency.

With many systems, this interdependency raises several problems for Critical Infrastructures, especially with the analysis or modelling of the infrastructure. Some of these factors, which could affect the interdependencies, were indicated by Rinaldi *et al*., such as timescales, geographic scales, cascading and higher order effects, social/psychological elements, operational procedures

business policies, restoration and recovery procedures, government regulatory, legal, policy regimes and finally stakeholder concerns [6]. These factors are critical and can have a negative impact on the system. Therefore, it is essential to understand some of the different modelling methods used in order to discover infrastructure interdependency [64].

One of the important pieces of evidence that has shown the interrelationship cascading possibility between infrastructures, was in 2001 when a chemical freight train derailed in Baltimore's Howard Street Tunnel [65]. Significant unexpected failures were found in different infrastructures, such as water, electricity and telecommunication systems.

2.3.1 Interconnectivity Modelling

Modelling individual infrastructures is a well-researched area, however, modelling of multiple interdependent infrastructures is still at an immature phase [6], [65]. However, the incidents of both natural disasters and cyber-attacks have proven that it is possible to measure the effect across infrastructures and find relationships through different interdependency modelling practices [66].

Rinaldi *et al*., grouped interdependency models into six different broad categories ranging from highly aggregated tools to very detailed, high-resolution and high-fidelity models [6]. The first category is the aggregate supply and demand tools category, which evaluates the total demand for infrastructure services in a region and the ability to supply those services. The second category is the dynamic simulation category, which can be used to examine a Critical Infrastructure's operations, the effects of disruptions and the associated downstream consequences.

In addition, dynamic simulation can be used to examine the effects of law, policies and regulations upon the operation of a Critical Infrastructure. The third category is that of agent-based models, which are used in a wide spectrum of interdependency and infrastructure analyses. The fourth

category is the physical-based model category where physical aspects of infrastructures can be analysed with some standard engineering techniques. This category of modelling can produce highly detailed information on the operational state of infrastructures, even down to the level of individual components. The fifth category covers population mobility models, which examine the movement of entities through urban regions as entities interact with each other. The sixth category is the Leontief input-output model category, which can be applied to infrastructure studies. This model provides a linear, aggregated, time-independent analysis of the generation, flow and consumption of various commodities among infrastructure sectors [6].

As an example of the modelling and simulation method, Di Giorgio *et al*., use a Dynamic Bayesian Network (DBN) as a novel technique for modelling Critical Infrastructure interdependency [67]. The DBN is divided into three levels. The atomic events level is concerned with using random variables that are linked to some unpleasant events reflecting on the Critical Infrastructure. The propagation level then constructs the frame of the interdependency in the Critical Infrastructure and their services. Finally, the services show the final Critical Infrastructure interest depending on the people working in Critical Infrastructure. However, Di Giorgio *et al*., faced some limitations particularly by using a specific discrete random variable and by limiting their failure types [67]. It has been considered that developing a new modelling method for the Critical Infrastructure interdependency is a challenge.

A second example is Interpretive Structural Modelling (ISM). According to Han *et al*., the ISM methodology can analyse the interactions of several Critical Infrastructures according to their mutual influence within a complex system [68]. Han *et al.,* applied ISM in their research on a system of eight infrastructures to develop a framework that shows interrelationships of Critical Infrastructures and classifies the different infrastructures' criticality according to their dependence

and power. Relationships that can be used to lead the whole system to be a more efficient infrastructure system were found as a result. However, these methods only show the relationship between the different systems involved.

By using both modelling and Critical Infrastructure interdependency simulation methods for data construction, to find the relationship between different Critical Infrastructure systems, this research aims to be able to reduce the impact of the cascading failure occurring to other infrastructures.

2.4 Cyber- Attacks

Companies around the world are investing possible solutions for their cyber-security concern, in order to protect their position and financial situation. CIP has become a prominent topic that can help in addressing cyber security threats [69]. This section presents, in depth, the cyber-threats facing Critical Infrastructures.

2.4.1 Cyber-Attacks and Critical Infrastructures

A significant number of Critical Infrastructures rely on the Internet. As a result, a number of security challenges have emerged, such as cyber-attacks and malicious activities [70]. Hence, cyber-attacks have become an important topic for researchers, military operations and businesses [71].

Various complex attacks are devoted to targeted Critical Infrastructures. On the whole, Critical Infrastructures have weak security systems and can be at risk of losing sensitive information, or Man-in-the-Middle attacks [70]. Chen *et al.*, for example, indicate that there was a 17-fold increase from 2009 to 2011 of cyber-attacks in the USA on the power grid infrastructure [72].

Abuzahkar *et al*., point out three cyber risks. These are: cyber-crime (managed by individuals or groups for extracting valuable information), cyber-war (this is usually conducted by nation states to steal valuable information or cause disruption) and cyber terror (this is managed by either an individual or organization working for a national state or for their own political views) [70]. In a successful cyber-attack scenario, the outcomes may compromise either the actual physical IT system itself or result in the loss of confidentiality, integrity and availability of an IT system [73].

Command *et al*., demonstrate several of the cyber tools that can be used in order to conduct their needs, which are listed as follows [15]:

- Backdoor
- Denial of Service
- Spoofing
- Keylogger
- Logic bomb
- Sniffer
- Physical Attacks
- Trojan Horse
- Viruses
- Worms
- Zombie

2.4.2 Infrastructures Security Threats

To date, successful cyber-attacks have either destroyed, manipulated information or deactivated stations [74]. One of these examples has been indicated by Wanger *et al.,* when a cyber-attacker managed to control and shut down the Prykarpattyaoblenergo Control Center (PCC) in the Ivano-Frankivsk region of Western Ukraine, which rendered more than 230,000 homes without electricity for up to six hours. This intrusion was one of the first strongly and well-organised attacks that happened against a nation's power grid [75].

Now taking the German nuclear power plant cyber-attack as an example, two viruses W32.Ramnit and Conficker were injected into PCs in the nuclear site. The viruses did not accomplish their target since the station was not connected to the Internet. These viruses were modelled in order to

snip and capture files from infected machines. Moreover, these viruses gave the ability to the attacker to control the system remotely with an internet connection [76] [77].

The smart grid is considered to be one of the most vulnerable sectors. Four common types of cyber-attacks can be found within the smart grid infrastructure [78]. Firstly, device attack, this attack usually starts at the initial step of an attack and focuses on controlling a grid device. Secondly, a data attack which is usually concerned within network traffic, manipulates data in order to alter decisions and results. Thirdly, a privacy attack which focuses on revealing private information from the electricity usage data. Finally, a network availability attack, which aims to delay or disable the communications in the smart grid by overloading the control centre with false information [79] [80].

All these examples show how vulnerable Critical Infrastructures are and how important it is to increase the level of security.

2.5 **Summary**

Critical Infrastructure assets contribute to the economy and society as a whole. Their impact on the security, economy and health sector are extremely vital. However, their increasing complexity has led to the creation of direct and indirect interdependent connections amongst the infrastructure groupings. This has resulted in different security threats facing interconnected security systems. By understanding the complexity of Critical Infrastructure interdependency, and how to take advantage of it in order to minimize the cascading problem, enables the prediction of potential problems before they happen.

Cyber-attacks are a major threat to the network. Critical Infrastructures will always be a significant target for cyber-attacks, as long as the Internet and the technology needed is available. Yet, cyber

security levels could improve if more communication and collaboration were found between the operator in private and public spheres. Shared information can limit the impact, help understand the threats and eradicate vulnerabilities.

CHAPTER 3

LITERATURE REVIEW

3.1 Introduction

The observation of physical data approaches towards Critical Infrastructures security focus on using network data. However, our system uses an observer pattern that recognizes threats and security breaches using behavioural observation and big data analytics.

Therefore, this chapter is a critical analysis of published sources. The chapter is formed to provide information about different facets that underpin Critical Infrastructures. These areas include big data and the use of analytics for predictive, descriptive, diagnosis and prescriptive analysis. Next, the chapter describes the publish-subscribe design pattern. Finally, a review of the machine learning is presented.

3.2 Big data

Over the past two decades, big data has become increasingly important in both the academic and the business communities [81]. In addition, big data has highlighted the need for infrastructures and tools to capture, store, analyse and use various amounts of structured and unstructured data [82]. The amount of data that passes through the internet represents a huge challenge for software developers and infrastructure companies in terms of how to operate, function and analyse this

volume of data, keeping in mind, this data needs to be securely delivered [83]. Therefore, companies have given invested considerable time to big data and the big data analytics in order to define data sets and discover the inherent patterns and anomalies that reside within it [82].

In order to cover the important points regarding big data, this section is divided into three subsections, starting with a brief discussion on big data and its associated characteristics. This is followed by a discussion on the challenges that face big data. Finally, a discussion on the four different analytic types is discussed.

3.2.1 Brief of Big data

Companies nowadays use big data to understand customer's behaviour and preferences. Big data is the next generation of computing that creates value through data scanning and analysis. Over time, data produced by users has grown exponentially for several reasons, including procurement data in supermarkets, commercial markets, banks, health and social networks. However, companies need to handle large volumes of data to control overload and noise. A number of definitions can be found to define big data. Jagannathan define big data as sets of large or complex structured and unstructured data that is collected from various sources [82]. In addition, Mishra define big data as a large amount of data that needs new technologies and architectures in order to extract value from it [84]. Large data consists of structured information, which accounts for 10% of the total amount of data available compared with the unstructured information that constitutes the rest [83]. Big data has penetrated into many communities, such as health infrastructures, governments infrastructures, and smart power grids [81]. Amazon is another example which computes millions of background processes every day and relies heavily on Linux systems to

handle the huge amounts of data generated from online orders. In addition, Facebook processes 50 billion user generated pictures [85].

Tole presented big data characteristics as the "3V's model", volume, velocity and variety. Volume refers to the throughput of data, velocity as the speed and frequency of data and variety as the different types of data and associated complexity [83]. However, Swapnil *et al.,* added two additional "V' models", which are: veracity which refers to the quality of the capture data and value described as the quality of stored data [86]. Figure 1 illustrates the multi V models.

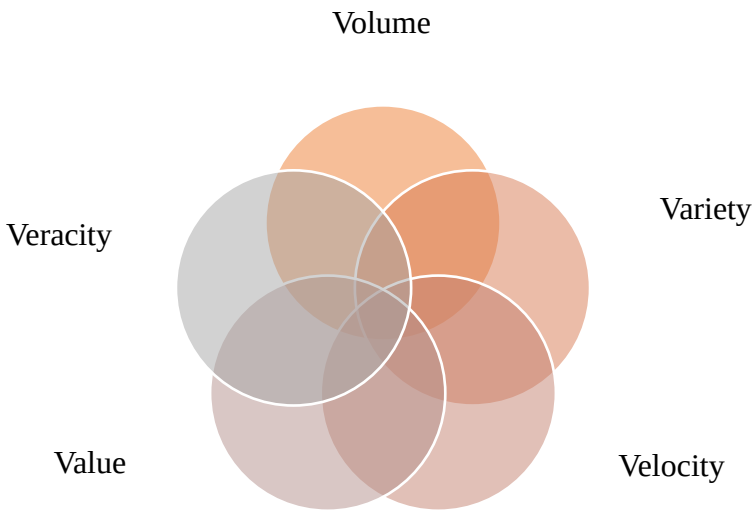


Figure 1 The Big Data Characteristics Models

3.2.2 Big Data and Analytics

Companies think that having data allows them to develop profitable results. The secret does not lie in data collection but in the relevant data you're using. Therefore, it is essential to determine the quality of data you need to compile in order to achieve a successful trade. Big data analytics is a method that helps organizations with decisions using large dataset to reveal hidden patterns,

anonymous correlations, statistical trends, customer behaviour and information that is valuable [87].

The analytics of large data seeks to improve decision-making and review data management decisions. Large companies use big data analytics system to improve the internal processes, such as risk management, improve existing products and services or to develop new products and service offerings [88]. Therefore, the benefit of big data analytics allows new opportunities to be uncovered. Big data analytics also helps to improve business decision making and helps to avoid crimes [87], [89].

Understanding the different types of big data analytics methods helps data scientists work effectively with their data. Therefore, four types of big data analytics are indicated by Figure 2.

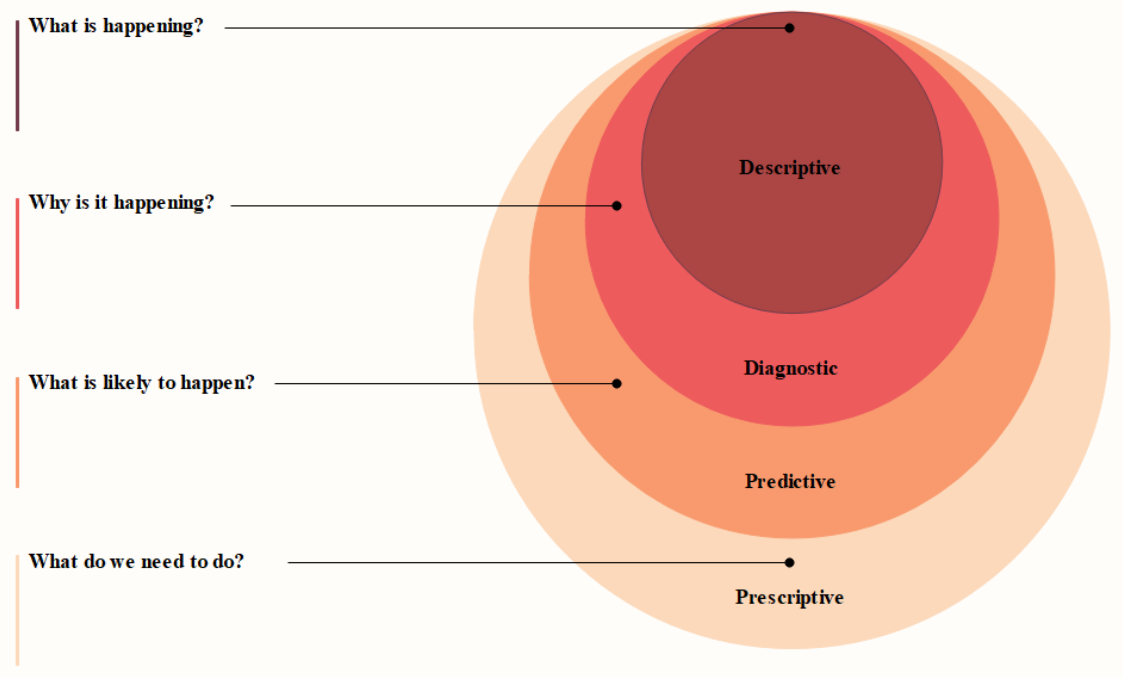


Figure 2 Analytic Types

Moreover, Pyne *et al*., and EMC describes these analytics types in depth [90], [91]:

- Descriptive: this method is the simplest class of analytics. This technique assists in understanding what happened historically and how this might affect future results. An example of descriptive analytics is credit risk and sales cycle.

- Diagnostic: this technique investigates why something has happened. In another word, this analytics method will discover the root-cause of the problem and isolate all confounding information. Moreover, the diagnostic technique focuses on the relationships and orders in the systems. An example of diagnostic analytics is a customer's health score

- Predictive: this technique helps in shaping the next step by answering, "What is likely to happen". This method depends on finding a correlation between variables, such as the age and heart attack risk. Keeping in mind that this technique uses historical data in order to illustrate the trends and patterns by using statistical models.

- Prescriptive: The purpose of this method is to make the best decision from the data adding some recommendation on why and how to make it happen. There are two approaches in order to present this technique: simulation and optimization.

Both descriptive and diagnostic analytics focus on the past and are used to summarize what happened while predictive and prescriptive deal with the future and summarize what can happen and how to make it happen [92].

There is a difference between analytics and analysis, the first is concerned with prediction and recommendation. While analysis is simply based on mathematical analysis [89].

3.3 Communication Patterns

This section presents an example of a communication pattern that has been adapted for the framework proposed in this Thesis, the Publish-Subscribe design pattern.

Publish-Subscribe is a method were the publisher aggregates the message to classes without the knowing who the subscribers are. Subscribers receive message events without knowledge of publishers. The objective of this system is to let information propagate from the publisher to the subscriber in an anonymous fashion. This type of message pattern offers great network scalability. Depending on the meta-data, publishers and subscribers share information cached locally. The publish-subscribe design pattern consists of three parts:

- Publisher: A publisher is an entity that creates notification messages based on situations. They provide information by creating a stream of messages that contain a header and a payload.

- Subscriber: A subscriber is an entity that receives notification messages from a notification producer. Subscribers are usually interesting in events.

- Message broker or event bus, is an entity that is responsible for registering to receive events from publishers and deliver events to subscribers. The event bus act as a temporary store.

This technique has the advantage of loose coupling, which allows the publisher and subscriber to discover and communicate with each other. Scalability in this thesis is one of the research challenges. Using the publish-subscribe pattern solves this. The design does not have an inherent impact on code performance. However, the publish-subscribe pattern does face delivery issues in

terms of time. Event loops can occur when there are many events published and subscribed to, and this can freeze the system. Figure 3 and 4 describe the process of the publish-subscribe pattern and how it has been adapted within our framework for communication.

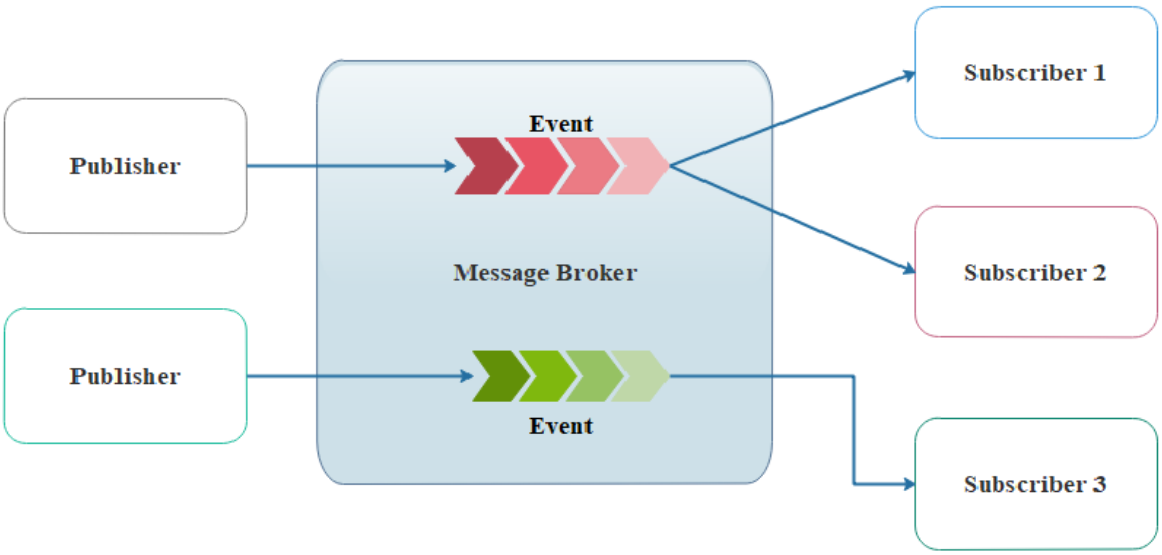


Figure 3 Publisher-Subscriber Pattern Process

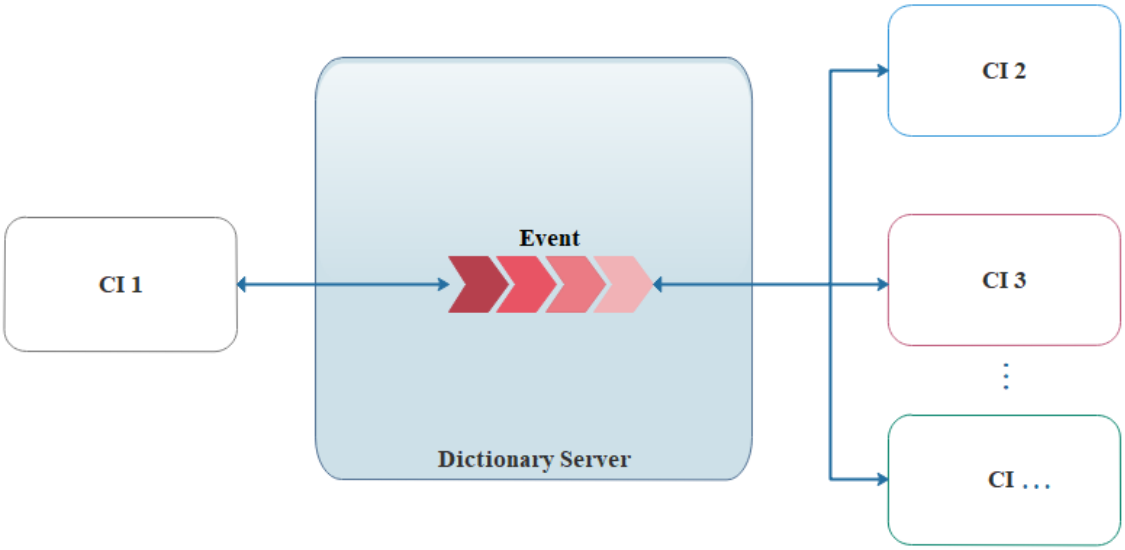


Figure 4 Adapted Pattern used in the Proposed Framework in this Thesis for Communication Processes

The figures shows the difference between the publish-subscribe communication pattern and the process for system communication used by the proposed framework in this thesis.

3.4 Simulation

Simulation has become an essential and useful part of mathematical modelling for many natural systems in physics, astrophysics, chemistry, biology, economics, psychology, social sciences and new technology architectures [93]. Simulation is an imitation process of real circumstances, which generally include some physical behaviours of systems, [94]. In recent years, the interest in simulation has increased for the purposes of education and experimentation. Typically, it involves the process of concepts, activities or experiments conducted through the computer [95]. Al-essa *et al*., define simulation as a method for teaching students that brings elements from the real world; overriding difficulties, such as, material cost or human resources [96].

Currently, there are various existing simulation programs, which contain '*smart*' built-in models for many common real systems. These programs help to analyse the inputs and outputs and do all the '*hard*' work required and give effective and comprehensive results.

Critical Infrastructure experimentation would ordinarily require the purchase of off-the-shelf hardware, which is extremely expensive and impractical for the average researcher [97]. This has led to the development of specific software-based simulators, such as Tecnomatix [98], and the adaptation of existing software-based simulators such as OMNET++, Simulink and Matlab [99]. These software simulators allow for affordable representations of critical systems, by modelling their behaviour, interactions and the integration of their specific protocols such as Modbus [100].

Generally, simulation can be divided into four types as Al-esaa *et al*., present [96]:

- Physical Simulation - This type of simulation uses the physical material to represent a testbed. This could be, for example, a flight simulator used for training pilots.

- Procedural Simulation - This type of simulation is designed to teach users a series of acts or steps; such as training on the steps to run a machine or device, or diagnosis of certain diseases in medicine. It allows people to train in a realistic environment without a real-world impact.

- Situational Simulation - This type differs from the procedural simulation, where the role of the learner is to discover the positions of appropriate responses by repeating the simulation.

- Process Simulation - This type of simulation helps to simulate the difficult movements, for example, detailed experiments, such as the movement of electrons. Therefore, the process of simulation can make it easier to grasp such concepts [101].

For simplicity, Al-esaa *et al*., divides the above four types into two main categories; either an educational simulation, through a hands-on experience, or through a visual demonstration by watching someone else [102].

Modelling and examining the interdependencies between Critical Infrastructures is considered to be a new essential field. Therefore, a significant amount of research is conducted into developing and modelling the Critical Infrastructure behaviour [103]. This is used by companies, government and communities in order to improve the performance of the traffic flow, manage the expenses, improve the reaction to emergencies [104]. In the following section, a discussion is presented on the benefits of simulation.

3.5 Why Simulation

The power grid was established long before the cyber-threat appeared. Although physical attacks are serious threats, cyber-attacks, nowadays, present a greater threat. Hence, current connected systems have a huge task to ensure that every individual component is checked and monitored [105].

Nearly all the European Union (EU) Member States have recognised the challenges of CIP and have advanced measures now in place [68]. Yet, it is still an ongoing and persistent challenge. Various factors contribute to this, for example, there is a lack of understanding about the interdependency scheme within Critical Infrastructure groupings. Moreover, there is no single

approach about how the elements of a Critical Infrastructure's functionality affect a connected partner [106], [102].

With many system types, this raises several problems; particularly with the analysing or modelling of infrastructure networks, which are relied on by multiple Critical Infrastructures. Some of these factors, which could affect interdependencies, are indicated by Rinaldi *et al*., [106]. They can include elements, such as timescales, geographic scales, cascading and higher order effects, social/psychological elements, operational procedures business policies, restoration and recovery procedures, government regulatory, legal, policy regimes and finally stakeholder concerns. These factors are critical and can have a detrimental impact on the system.

As previously discussed, most Critical Infrastructures are controlled by SCADA, which is one of the infrastructure challenges that make it vulnerable to cyber-attacks [107]. Linking the SCADA network to business networks has increased threats, such as malware, insider, hacker and terrorism [108]. SCADA attack consequences are particularly serious, affect the public and cause a financial crisis. An example of this was seen in the Aurora Generator Test, carried out at the Idaho National Laboratory in 2007, which cost the US $1 million dollars by simulating a remote cyber-attack on a generator control station [109].

Therefore, simulation has a key role in the advancement of CIP. Its use is becoming a common technique for the testing of cyber-attack prevention measures and for improving the level of security techniques [110]. A simple system can be created to represent a larger infrastructure and allow for realistic testing to take place. It is clear that there are many benefits of using simulation. It helps in conducting experimentation on a realistic representation of a system, without the worry that any damage that was done would have a real impact [111]. In particular, when testing against

cyber-attack resilience and developing new approaches to security, Critical Infrastructure simulation is of great benefit [40].

Ündeger *et al*., further discuss that modelling and simulation offer time-efficiency [41]. This is particularly beneficial for understanding phenomenon-based disasters and exploring the impact of new policies, without testing it in the real world. Ündeger *et al.*, recognise that using simulation approaches to model interdependency within a complex system can answer the major '*what-if*' questions within a safe environment [112].

To date, simulations have helped in developing and enhancing new framework concepts to improve security levels. For example, NIST developed a simulation framework to reduce the risk of cyber-attacks to Critical Infrastructures [113]. The NIST framework includes sets of procedures and methodologies that help to understand the cyber risks. Moreover, the approach involves flexible, classified, performance-based and cost-effective methods with more security measures. Specifically, the NIST cyber-security framework has been set up to strengthen security through the following:

- Diagnose the security status of a system.
- Mend and form a cyber-security program.
- Detect new chances for new or known standards.
- Support Critical Infrastructure organisation, to use a cyber-security framework with tools and technologies [42].

As such, Critical Infrastructures have benefited from the NIST cyber-security framework. This has been recognised by some notable improvements, such as reducing the time of starting the security program, recognising the improving areas in the program and improving the efficiency with law communication between Critical Infrastructures [113].

As discussed, simulation can be employed to help manage the understanding of interdependency between Critical Infrastructures. As such, the next sub-section presents the different modelling work that has been undertaken in order improve the level of security currently in place.

3.6 Critical Infrastructures Modelling

Several works have investigated cascading failures in Critical Infrastructures, and subsequent models that can be used to help simulate the impacts of failure to improve the level of security. Therefore, this section presents a number of modelling and simulation examples.

Table 2 and 3 present a summary of the different models that were used in order to enhance the impact caused by cascading failures in Critical Infrastructures. Laprie *et al*., for example, choose electricity and the associated information infrastructures because of the massive growth in their interdependency, which raises the vulnerability that can occur from the cascading and increased risk of large scale blackouts from the electricity transmission system. Laprie *et al*., used a qualitative model focused on the main failures that happened from interdependency such as, cascading and expanding [114].

Table 2: Summary of Studies on Infrastructure Modelling and Simulation Examples

Authors	Description	Critical Infrastructure	Findings/ Outcomes	Limitation /Gaps	Our Model (CIAIRS)
Haimes *et al.,*	Develop a methodology that administrates the risk of cyber-attacks on interdependent infrastructures especially that target the production, filtering and supplying. The problem in this system is that it is highly interconnected and interdependent with the inner system, the external infrastructures system and the financial sector.	Oil & Gas	Develop a better understanding of the probability of the cyber-attack on the measures, infrastructure and economic level. Indicate the level of risk that can be posed to SCADA from the infrastructure interdependency. Recommendations for managing cyber-attacks in interdependency infrastructures	• According to the massive system area that accrues a problem in the modelling dimension. • As a result of the limitation in money, the model needed a risk identification phase to structure the critical scenarios. • The system needs to be built on existing models and analyses	Built on existing models and data that helped in building a correct form.
Byres *et al.,*	The model that was used focused on the communication in SCADA, based on Modbus protocol. The methodology was created depending on the	SCADA communication	Indicates a number of features for 11 attacker goals, such as the attacks goals, the level of the attack difficulty on SCADA, the impact and the likelihood of these attacks and the probability of spotting the	Did not consider the attacks from: other interdependency points, SCADA communication by other infrastructures, the wireless connection and third parties policies	Covered the interdependency between the infrastructures and studied the effect of this connection and how it can help in alarming the

	possibilities and the interconnectivity between the attacks. 11 attacker goals were used to improve the Modbus standard		attacks. The attacker goals were mainly focusing on identifying the Modbus and gaining access to the network.		other infrastructures
Laprie *et al.*,	Present the failure related to two Critical Infrastructures from their interdependency by a qualitative modelling. The model was a qualitative structure depending on the behaviour of the interdependency between the two infrastructures	Electricity and associated information	Different models were found and refine the impact of the quantitative measures on the interdependency between the electricity and information infrastructures.	Focusing on only two types of attack deceptive and perceptible attacks and isolate the impact of accidental faults from the malicious fault is considered as a limitation of this work that made the model non-integrated	More attacks and more infrastructures taken into account in order to give a bigger picture of the effect of the interdependency within the infrastructures.

In another approach, researchers discussed the huge impact of blackouts on the economy caused by cascading failures, and how they use modelling to capture the dynamics of cascading failures and blackouts in electricity power systems [115] [116]. Dobson *et al*., discuss the CASCADE model, which focuses on the critical point and the stage where the cascade failure is likely to happen [117]. Finally, complex network theory is used to model cascading failures in infrastructures. This method depends on analysing the structure and flexibility of the network and removes accidental failures or attacks [118]. Therefore, Kinney *et. al*., use the Crucitti-Latora-Marchiori (CLM) model as one of the complex network models based on a dynamic approach [119]. Next Chassin *et. al*., employ the Barabási-Albert network model '*apple*' to generate scale-free networks using connected nodes that use the power-law connectivity probability [120].

The above models used complex network theory to model the cascading failures in electricity infrastructures to consider what will affect the system and direct the system to breakdown. However, researchers have used models that have not considered the complexity of the relation and the interdependency between the electricity infrastructure and information communications infrastructures. Finally, it is clear that the modelling did not cover well-known failures, such as escalating failures.

Table 3: Summary of Studies on Infrastructure Modelling and Simulation Examples

Authors	Description	Critical Infrastructure	Findings/ Outcomes	Limitation /Gaps
Dobson and Carreras	Probabilistic model created to spot the cascade in a system and clone new features to understand the failures from the blackout and cascading failures. Use the critical point in order to estimate the probability of cascade failures	electricity power	By shaping the model and operating it to reach the critical point where the system becomes in risk of cascade failure the blackout risk ranked between small, medium and high	It was hard to balance between the cost and limiting throughput in the system from a costly common cascading failure and management of blackout risk
Kinney *et. al.*,	Present the grid as an undirected graph and analyse features from the system flow as internet traffic and power flow. CLM present the power grid as nodes, generator, loads and connective edge presenting the transmitted lines	power grid	Helped in studying the power system characteristic at a comprehensive scale and assist the level of cascading within different network topology	CLM can face some limitation of measuring the blackout size

Chassin and Posee	Have analysed the reliability of North American eastern and western electric grids by using the complex networks model the Barabási-Albert network model to improve the grid flexibility. Two failures were identified that helped in cascading failure in the electricity grid: edge removal and node removal	power grid	Indicate that the relation between the node and the interconnected very efficient to indicate the behaviour in the scale-free systems. Produce a model indicating the cascade failures in two electricity grids by knowing both the physical and the policies in the system. merging the probability connection for the scale-free network to evaluate the reliability for the two energy infrastructures	The model used only two reliability models, which made the result of the eastern and western electricity grid not efficient.

3.7 Machine Learning

Machine learning allows computers to learn and discover insights in data without being programmed to do so. Machine learning improves computer functionality when exposed to new data and uses previous computation in order to produce dependable, repeatable conclusions and results. Various field use machine learning to recognize the data they have, such as financial services to prevent fraud, government utilities to reduce identity theft, and the oil and gas industries to improve oil distribution lines [121]. In machine learning, computers are not told how to solve a particular problem, instead algorithms are utilised to abstract meaning from data based on different learning models – supervised, unsupervised, semi-supervised and reinforcement.

Figure 5 and 6 indicates these methods in a simple clear way.

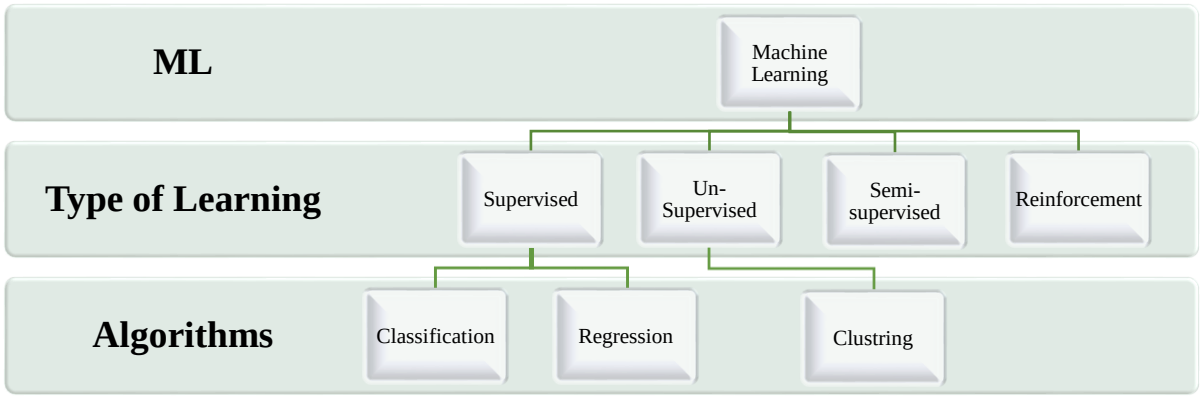


Figure 5 Machine Learning Algorithms by Learning Style

Algorithms are grouped by learning style: supervised learning and unsupervised learning are the two main types. However, there are other methods, such as semi-supervised and reinforcement. The learning algorithms focus on the roles of the input data and the process of choosing the model in order to select the most suitable to reach the best outcome. In supervised-learning algorithms,

historical data (training dataset) is used to train models. Supervised machine learning algorithms predict the values of labelled data depending on patterns learned patters within the data [122]. In the case of unsupervised-learning algorithms, unlabeled data is utilsed to cluster data into different groupings to try and find structures. This is often referred to as cluster analysis. Semi-supervised machine-learning is a mixture of both supervised and unsupervised, which deals with labelled and unlabelled data sets [123]. The dataset in this technique use a small amount of labelled data with a large amount of unlabelled data [124]. In the case of reinforcement-learning, algorithms focus on balance examination of unknown data set with the use of current knowledge. This technique does not need correct input-output data sets - it only needs to learn good policy [121].

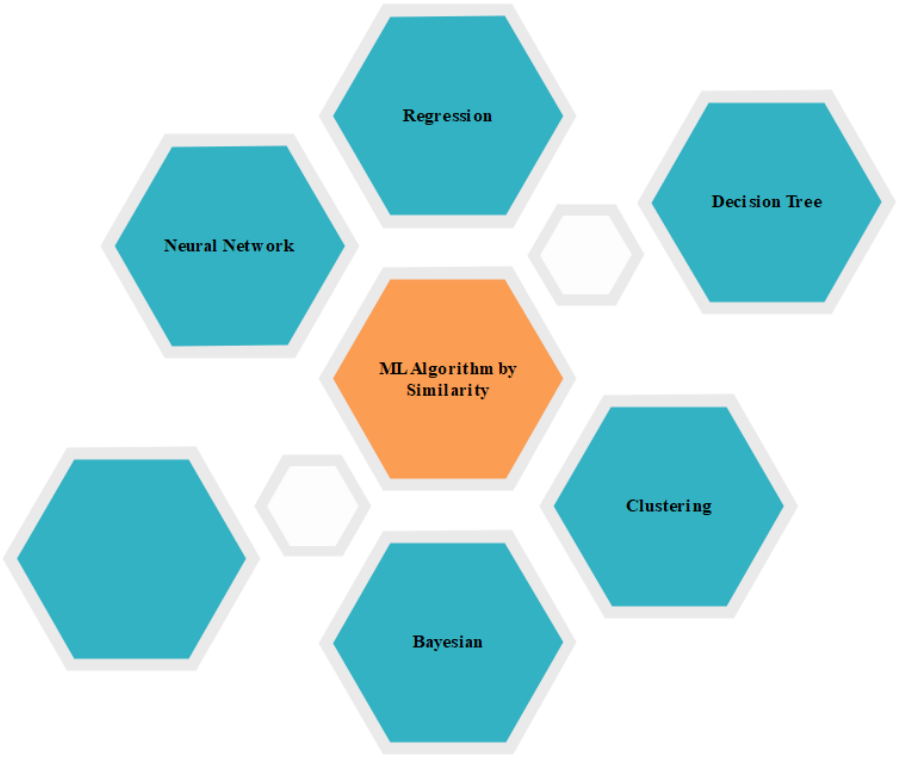


Figure 6 Machine Learning Algorithms by Similarity

3.7.1 Supervised

In this thesis, the evaluation uses supervised learning as illustrated in Figure 7. In this approach algorithms receive known input data called training data along with a known label output for each observation. This learning type provides a mapping function between input variables (x) and output variable (y). The goal is to predict output (y) from new input variable (x) [125]. Under supervised machine learning, there are two categories included: classification and regression. Classification is used when the output variable can be grouped into classes. An example of classification algorithms are support vector machines (SVM), neural networks, Naïve Bayes classifiers, decision trees, discriminant analysis, and *k*-nearest neighbour (kNN). Regression algorithms are used to predict a continuous value or a real output value. Common regression algorithm includes linear regression, nonlinear regression, generalized linear models, decision trees, and neural networks[122].

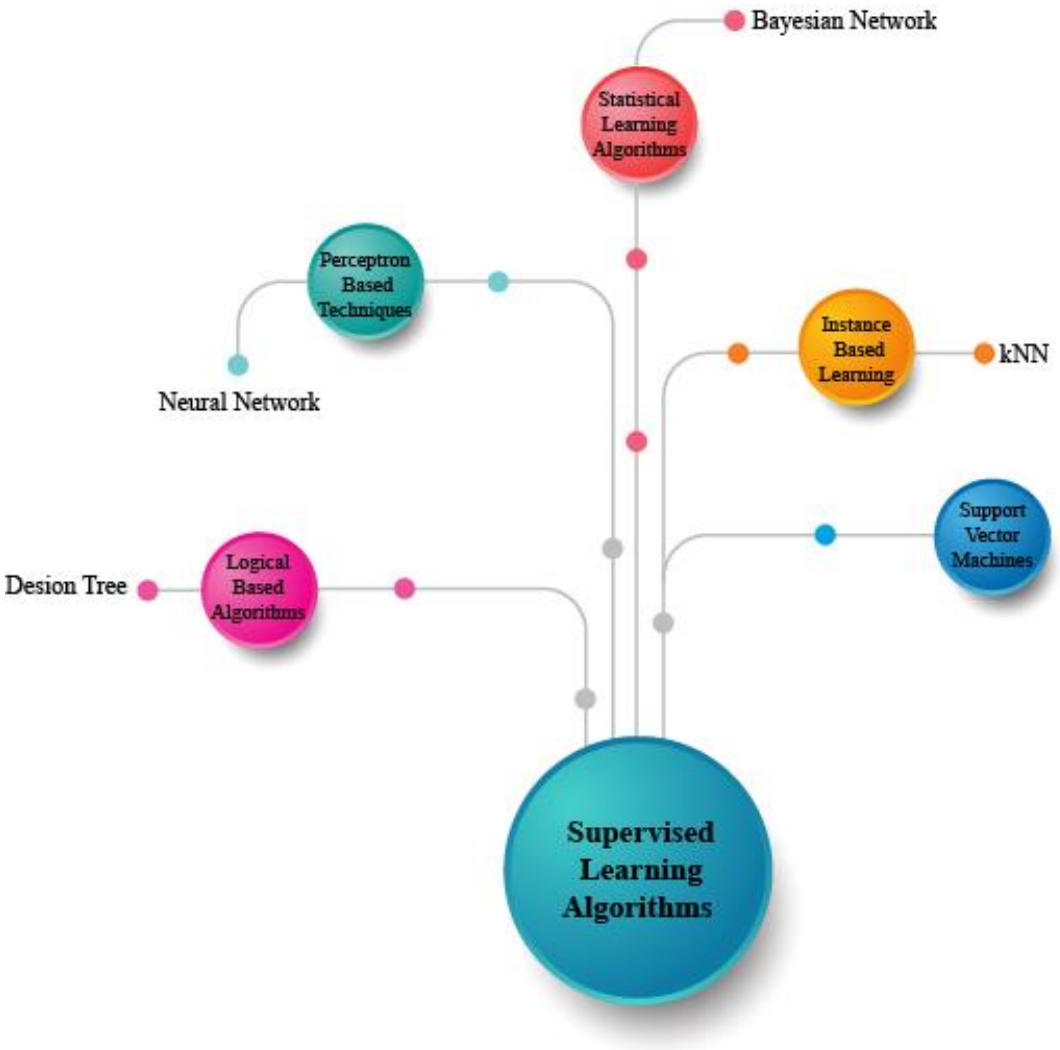


Figure 7 Supervised Machine Learning Techniques

Figure 7 describes five supervised learning algorithms. The first algorithm is logical based that use decision trees as an example. The decision tree is a non-parametric algorithm that classifies data depending on the feature values. This type of algorithm consists of nodes and branches that handle missing values and interactions between features. Each node represents a feature and each branch represents a value. Although the decision tree provides high performance, it is hard to deal with high dimensional data. Moreover, one of the serious problem that face this algorithm is error propagation through trees. Next, the perceptron is a well-known algorithm capable of dealing with

non-linear or dynamic data. A perceptron is the simplest form of Neural Network (NN) that is used to classify linearly separable data [126].

Bayesian Network classifiers, unlike NNs, take less computational time to train and there is no need for hyper parameter tuning. However, the size of the data can affect performance in Bayesian classifiers. Another algorithm that is considered a lazy-learning algorithm is the nearest neighbour algorithm (kNN), which is a non-parametric method. The algorithm is based on a distance function. kNN takes less time to train but more time during classification. The algorithm has a number of disadvantages, such as large storage requirements [127].

Finally, the support vector machine (SVM) algorithm finds the best hyperplane to separate data classes from each other using the largest margin. It makes a simple linear separation between classes. This algorithm offers high accuracy and performance is not effected by the size of the features. However, the complexity and the choice of tuning parameters can affect the performance [121].

Several considerations need to be made when choosing appropriate learning algorithms, such as memory usage, prediction rapidity and the model explanation. In addition, the predictor level is important to take in count whether continuous or categorical results are required [85]. The evaluation for any classifier depends on prediction accuracy. There are two techniques to calculate the accuracy of classifiers: the first technique is by splitting the training data and use two-thirds for training and the rest for estimating performance. Cross-validation is the other technique, which works by dividing the training set into mutually exclusive and equal size subsets. The classifiers estimate of error depends on calculating the error rate of each subset [122].

3.8 Summary

This chapter presented literature on big data and different analytics types: descriptive, diagnostic, predictive and prescriptive. Moreover, a brief description on classification approaches, was provided, grouped by supervised and the unsupervised learning algorithms. After understanding the strength and limitation of each method, few algorithms will be used to solve the problems investigated in this thesis.

Effective security is costly, especially as Critical Infrastructures often require security systems to be tailored to match their unique needs. Simulation can play a huge role in filling this gap. Different modelling examples can be found to show how modelling helped in raising the level of security by understanding cyber-attacks.

Table 4 indicates the different models, which the authors have used in order to simulate and model Critical Infrastructures. Haimes *et. al.*, for example, explain the use of the Hierarchical Holographic Model (HHM) and how it can assists to resolve the extensive system into a hierarchy of subsystems that give an overview of the system from different angles and capture the interdependency within the Critical Infrastructures [128][129]. The HHM was established to solve the problem of the single model ability and demonstrate each part of the system as individual subsystems to analyses the interaction between each other. Therefore, a number of points need to be available from the system as the role, activities, political restrictions and the structures of the system [130]. In another example, Byres *et. al.*, focus on using the tree methodology in order to model the cyber-attacks. The attack tree assists in spotting the damage to SCADA systems by identifying the flaws, produces a special test from measurable goals that can be matched to real world devices, networks and implemented. Moreover, the tree applied at multiple layers considering other interruption [131].

Table 4: Summary of Author's Different Models and Simulation Examples

Authors	Case Study	Simulation	System	HHM	Tree	Qualitative Structure	Complex Network Theory
Haimes *et al.*,	√	√		√			
Byres *et al.*,		√			√		
Laprie *et al.*,		√				√	
Dobson and Carreras		√	CASCADE				
Kinney *et al.*,		√	CLM				√
Chassin and Posee		√	Barabási-Albert network				√

CHAPTER 4

CRITICAL INFRASTRUCTURE AUTOMATED IMMUNO-RESPONSE SYSTEM (CIAIRS)

4.1 Introduction

By understanding the complexity of Critical Infrastructure interdependencies, potential problems can be predicted before they happen and cascade throughout the network. The system proposed in this thesis is able to detect cyber-attacks, share the knowledge with interconnected partners and communicate the potential impact by creating an *immune system network*. Our work developed a system called Critical Infrastructure Auto-Immune Response System (CIAIRS), which assists and guides Critical Infrastructures on how to behave when abnormal behaviour is detected. Anomalous behaviour is then shared with other infrastructures and has been inspired by the human immune system characteristic [132].

This chapter begins by; identify specific requirements that are important in designing the CIAIRS framework. Next, a number of real-life scenarios are highlighted in order to identify the system

requirements. This is followed by the CIAIRS structure along with a detailed account of the various components that work together to predict abnormal behaviours and share decisions with other infrastructures. Finally, the functioning of the system is detailed.

4.2 CIAIRS System

CIAIRS functionality relies on identifying attacks and guiding Critical Infrastructure operators on how to behave when abnormal behaviour is detected. Furthermore, inspired by the human immune system characteristic, the information is then shared with other infrastructures to create an artificial immune system network [132]. The quality of a framework depends on four main features: Simplicity, Clarity, Boundaries and Expandability [133]. Therefore, these features were taken in account when forming the research approach.

4.2.1 CIAIRS Framework Requirements

This stage identifies a number of requirements, which helped in designing the CIAIRS framework.

The requirements of the system are defined as follow:

- The CIAIRS framework has to be flexible and have the capability to add new features. Therefore, the system should act as a plug-in service

- The CIAIRS framework has to collect data for analysis and reporting. In addition, transforming and merging data functions are required.

- The new system has to be scalable and adaptable to different Critical Infrastructures.

- The new framework has to learn system behaviour.

- The system needs to identify system anomalies

- The new framework has to identify the interdependency between Critical Infrastructures.

- Depending on interconnectivity registration, the system has to share the results of abnormal behaviour with the rest of the Critical Infrastructures

- The system needs to use an operator to check the system activity

- The system can save responses from different Critical Infrastructures about abnormal behaviour.

4.2.2 CIAIRS Location

Depending on the type of Industrial Control Systems (ICS) that monitors and controls the industrial infrastructure, the location of CIAIRS resides between the network layer and the enterprise layer [134]. CIAIRS is a comprehensive plug-in superintendent package for collecting and sharing abnormal behaviour with the rest of the Critical Infrastructures. In other words, the CIAIRS is a connecting software component in the network that collects data. Consequently, from a security point of view, if the CIAIRS was breached or attacked, this will not affect the rest of the infrastructure, since an operator is responsible for making the final decision before sending the information to the rest of the Critical Infrastructures, as shown in Figure 17. In order to picture the CIAIRS location, the ICS layers will be presented then the CIAIRS location. The SCADA system includes four major features: the operator or human machine interface (HMI), master terminal unit (MTU), communications, and remote terminal unit (RTU). Figure 8 specifies the ICS layers and Figure 9 illustrates the CIAIRS position within the ICS layout and the Industrial Network (IN).

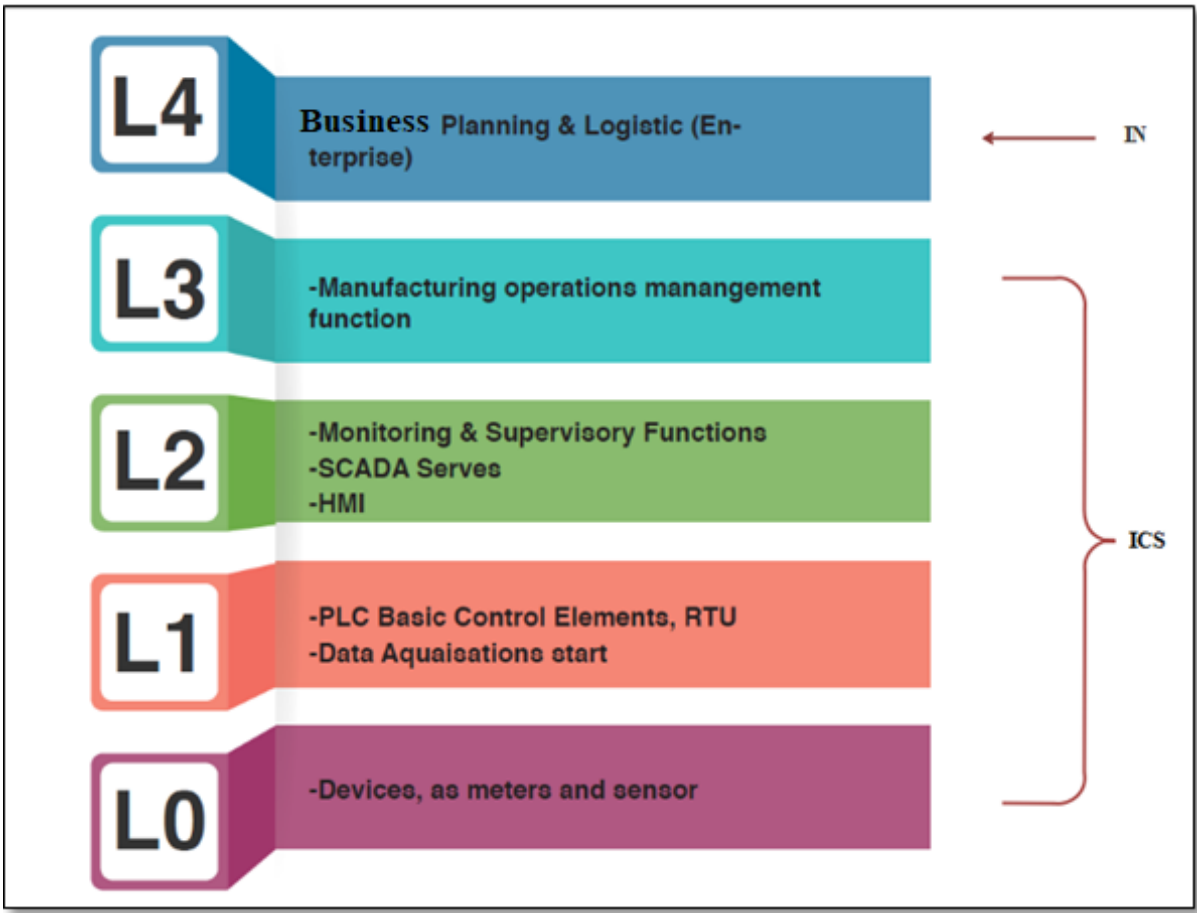


Figure 8 The Industrial Network Architecture

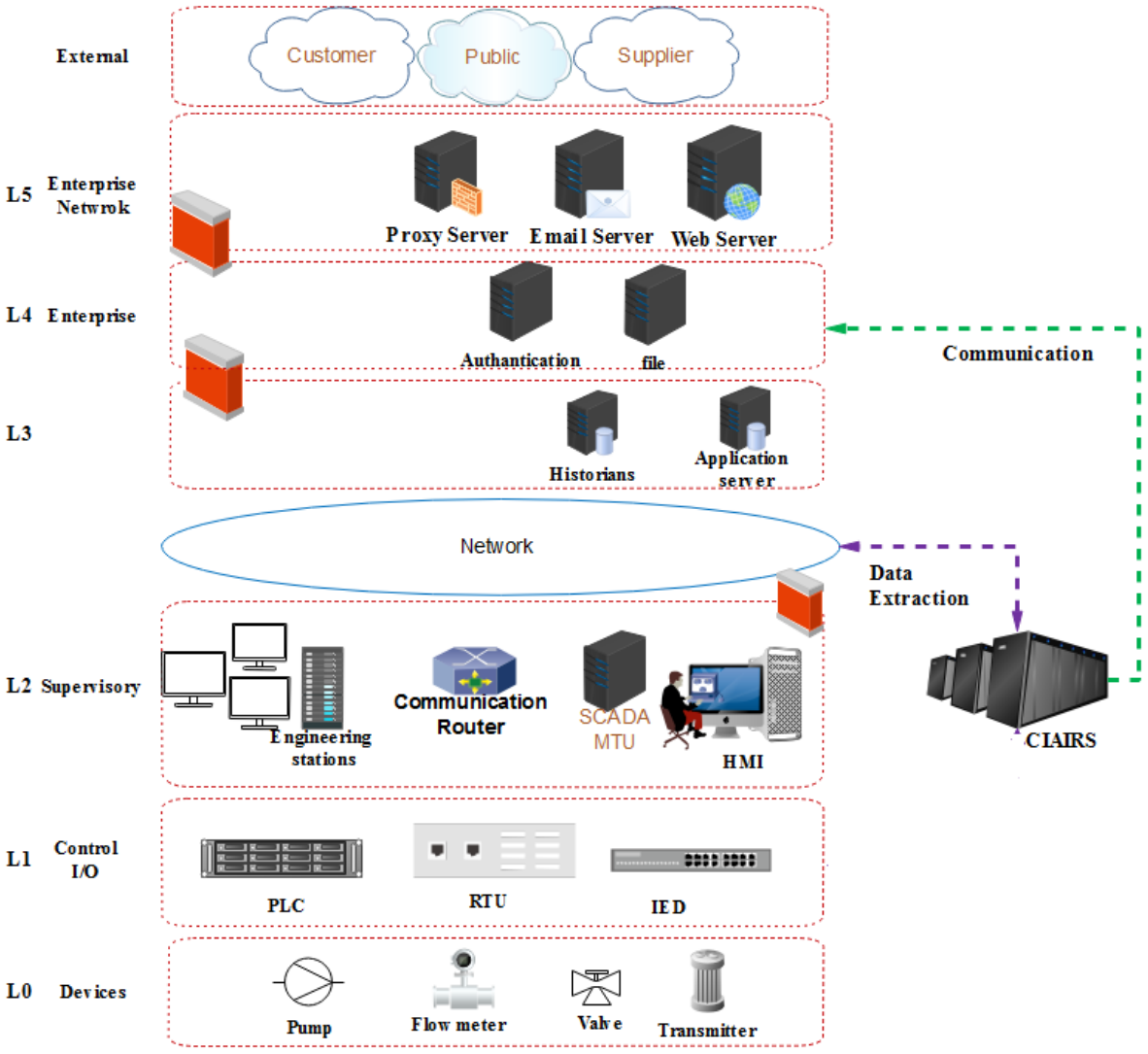


Figure 9 CIAIRS Location within the IN & the ICS layers

A number of applications are executed by CIAIRS, such as the Distribute time stamp, which assists in viewing the behaviours in the system; and numerical distribution, which helps convert numerical data to scales, such as high and low. Statistical export sheets, present the effect of an attack more easily than data view. This data is made use of by the CIAIRS system.

4.3 Real Scenarios

Four real situations are presented relating to a Water Critical Infrastructure located at the National Ministry of Water in Jeddah, Saudi Arabia [135]. These scenarios are listed as follows:

- Displacement in the main water carrier line from the source, which moved three sub water lines and affects other parts of the infrastructure.

- Hit the main water line distributor with a shovel while extending power cables. The damage appeared 10 days later.

- Separation in a water domestic line by a motor car, which caused water to leak for three hours.

- A contaminated water leak that occurred in the main water line which stopped the water pump and suction for the contaminated water.

The way to handle these scenarios was different depending on the amount of the damage in the water line, the equipment availability, the laboratory analysis results and the damage to other infrastructures.

4.4 The CIAIRS Framework Construction

Figure 10 illustrates the CIAIRS framework and the interaction between the various modules, which function together to perform the security and communication services. The modules linked together form the system as a whole, and work together in order to detect abnormal behaviours in one infrastructure and share them with others. The aim is to prevent cyber-attacks from having a

cascading impact and spreading to other infrastructures. Threat information is communicated to allow operators in other infrastructures to take appropriate measures to prevent an attack having an impact.

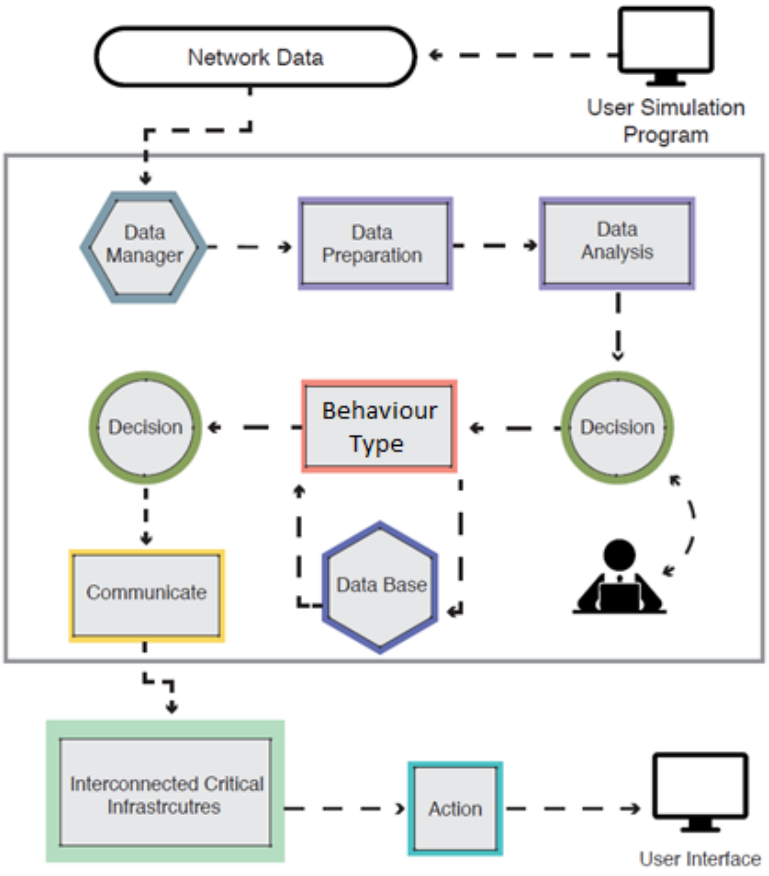


Figure 10 CIAIRS Construction

The whole process is clarified further in the next subsections, which detail the major key modules that comprise the CIAIRS system.

4.4.1 Data Manager

Within the Data Manager, the database holds different types of data as text, numbers and other readable forms. The data manager is responsible for directing the information collected from the

system to other modules. In addition, the data manager is responsible for all aspects of data processing, such as feature extraction, entry, validation, manipulation, backup, purifying and storage.

Figure 11 presents a Data Flow Diagram for CIAIRS data processing of infrastructure data. A Data Acquisition (DAQ) unit is used to store data blocks to prevent system overloading. The data is either stored in the Critical Infrastructure DB or sent to the next step, for preparation and feature engineering. Each feature extracted takes time (T) and has data status (S). The status varies between sleep, repair, slow, overload, normal and cyber. These different statuses allow the system to realize the difference between normal and abnormal behaviour.

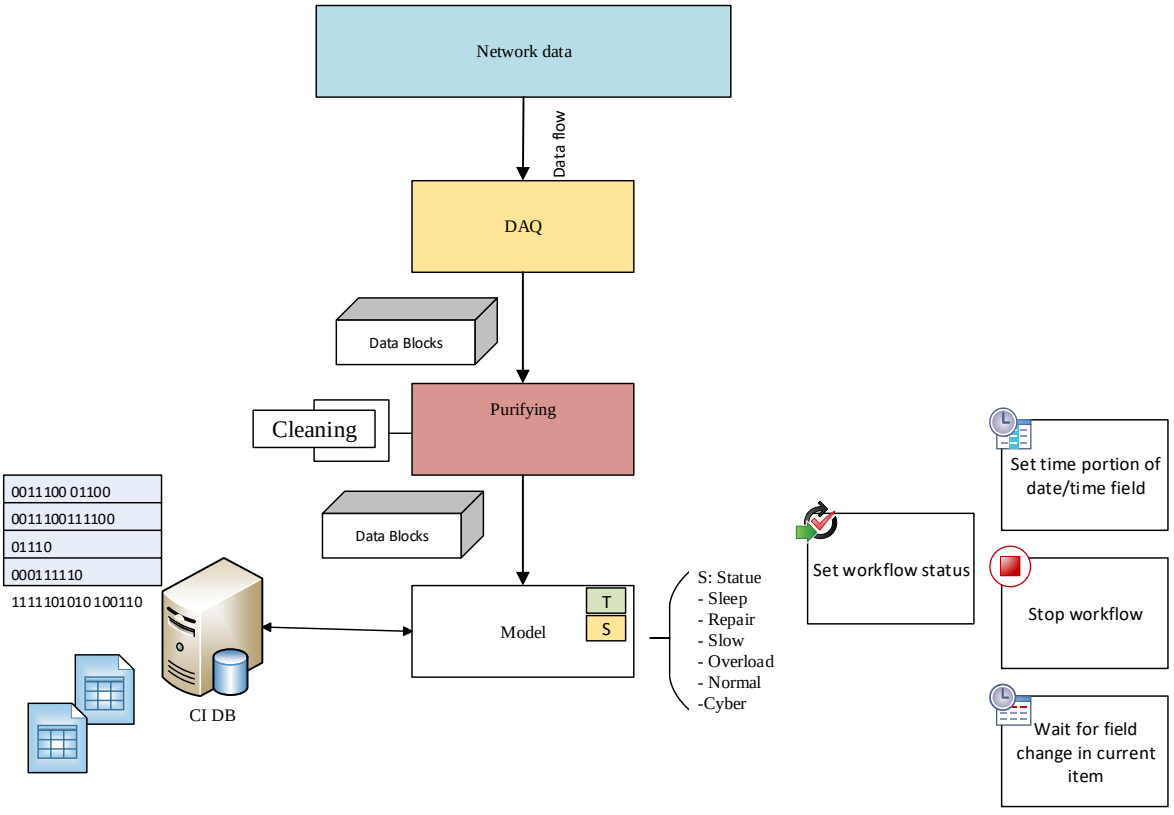


Figure 11 CIAIRS Data Manager

4.4.2 Data Preparation

Real world data is often incomplete, noisy and inconsistent. Therefore, a data preparation process is important in order to clean, integrate, transform and reduce the data collected from the network. The data is passed through a cleaning process to make sure there is no missing information, reduce the noise in data and eliminate inconsistencies, such as duplication values; this stage is called the Data Preparation stage. Next, the data is normalized to ensure the values follow a specific range, usually between 0 and 1. Consequently, the data is split into multiple fragments to reduce the redundancy and to avoid data anomalies. By normalizing the data, the database is more suitable for querying. At this point, features are extracted analysed. The features correspond to the system behaviour and present a simplified view of the overall network. Figure 12 displays these stages.

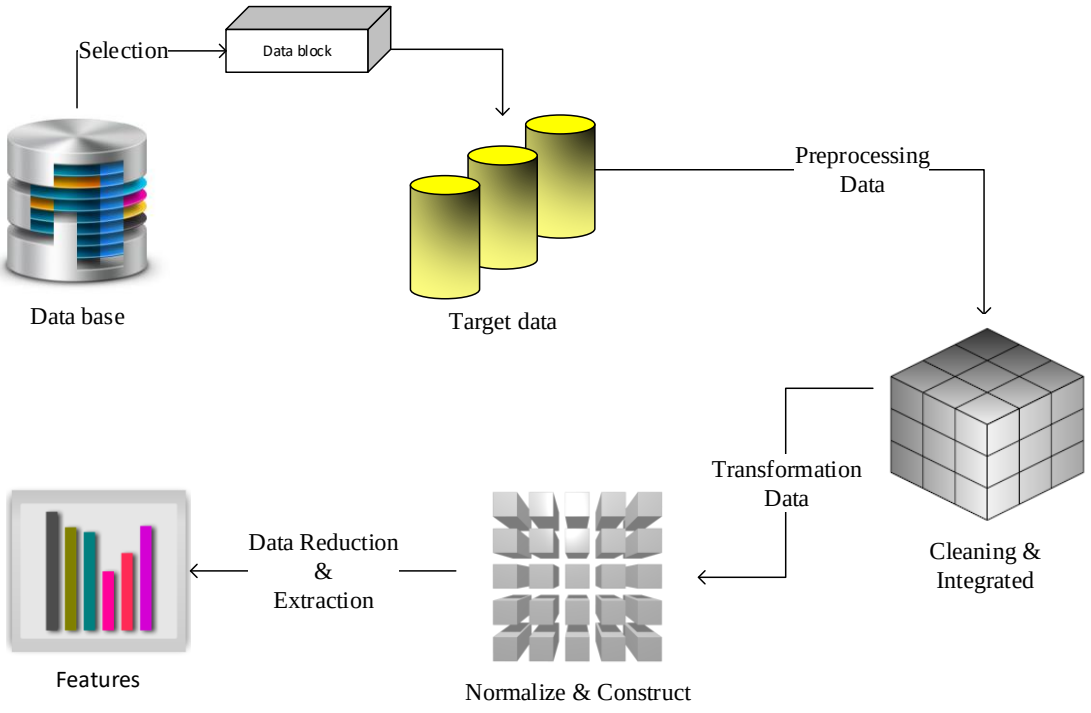


Figure 12 CIAIRS Data Preparation

4.4.3 Data Analysis

Using a feature set, classification used to model normal and the abnormal behaviour. Figures 13 and 14 illustrate the main function of the CIAIRS analysis. Figure 13 illustrates the feature extraction process, which depends on the Critical Infrastructure type.

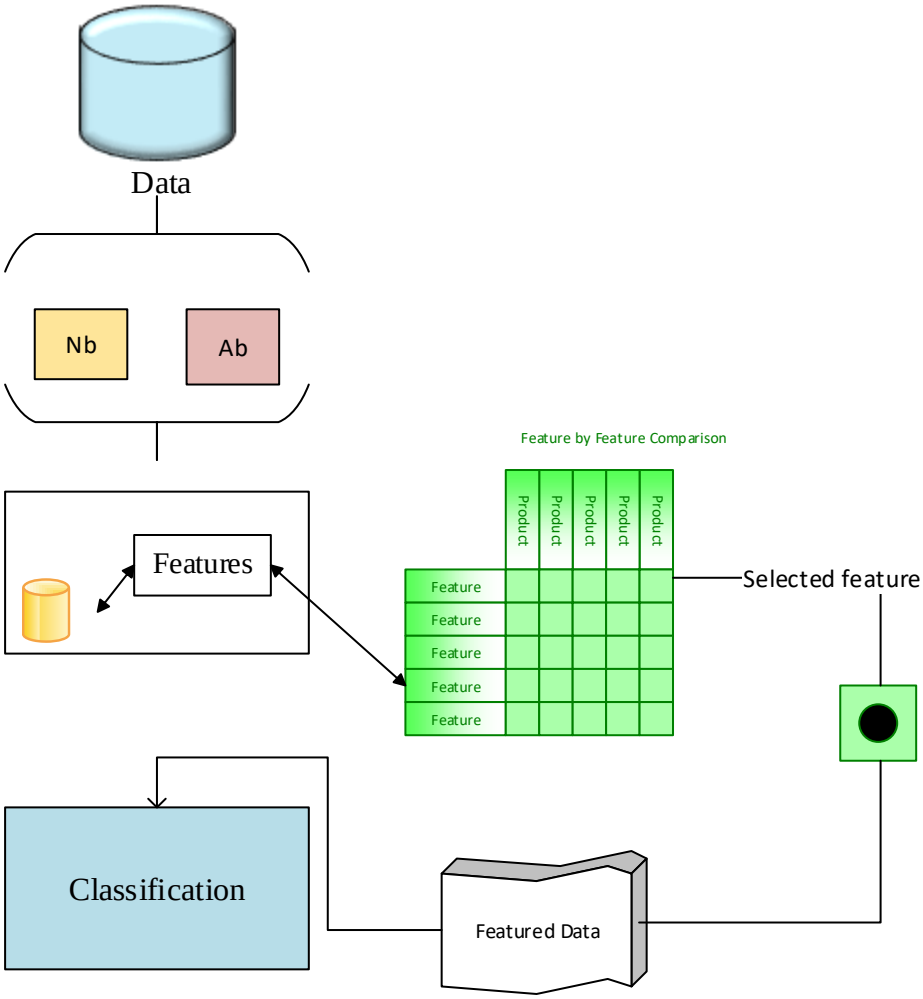


Figure 13 Feature Draw out

The goal of using these features is to discover useful information in the data and simplify classification modelling processes.

By classifying the features, the data can be identified as either normal or abnormal system behaviour. Normal behaviour (Nb) would be when the infrastructure is functioning under correct parameters. Whereas abnormal behaviour (Ab), would be when the system is functioning incorrectly as a result of a cyber-attack taking place. The classification process enables CIAIRS to distinguish between the two types. Figure 14 specifies the CIAIRS classification process, using the features to classify normal and abnormal behaviour.

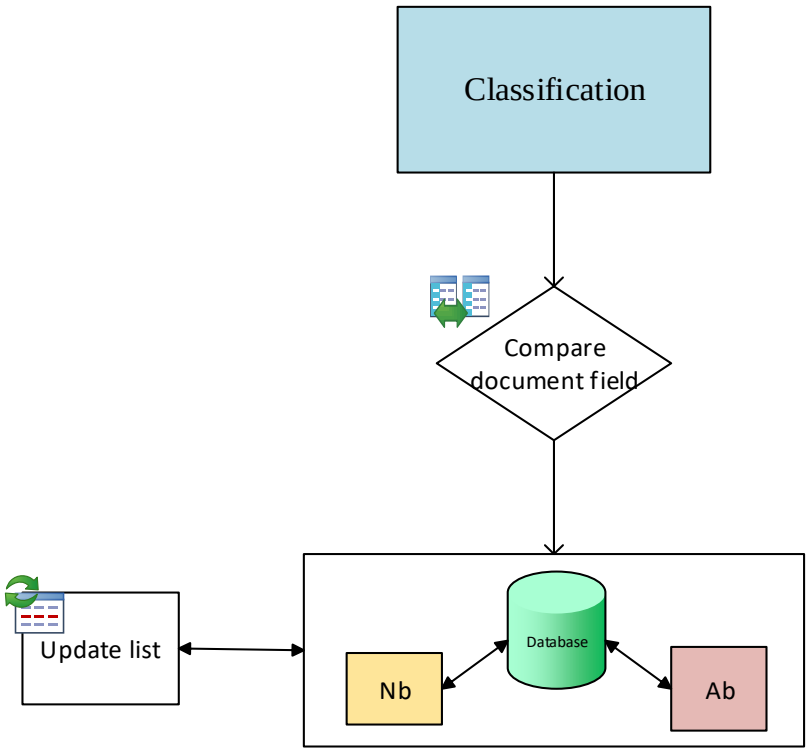


Figure 14 CIAIRS Classification

4.4.4 Attack Pattern Recognition

CIAIRS is not only able to detect abnormal behaviours in the system but, also based on the immune system concept, CIAIRS is able to recognise the behaviour of specific attack types. In order to do this, it employs a pattern recognition and comparison process. The recognition is derived from a database supervised by the operator.

4.4.5 Database

Five different databases are used in the system in order to hold the following information:

- The saved feature list
- The pattern of the abnormal and normal behaviours
- The ID for the different Critical Infrastructures in order to identify the interconnectivity
- The CIAIRS shared information and the known attack pattern.

Each of these databases uses SQL, which is a common database type employed in Critical Infrastructures.

4.4.6 Interconnectivity

Discovering the interdependency between the different Critical Infrastructures is considered one of the main features of this system. After identifying abnormal behaviour, and determining the attack type in one of the Critical Infrastructures, the attack information and characteristics are shared with interconnected partners in order to prevent cascading failures. In order to function correctly, a CIAIRS system would be needed in each partner's network. The system starts by sending a copy list that includes the new abnormal behaviours. However, in order to send it, an interconnectivity check is needed. Therefore, a communication server is needed to hold the different Critical Infrastructures ID. Figure 15 presents the interconnectivity check process.

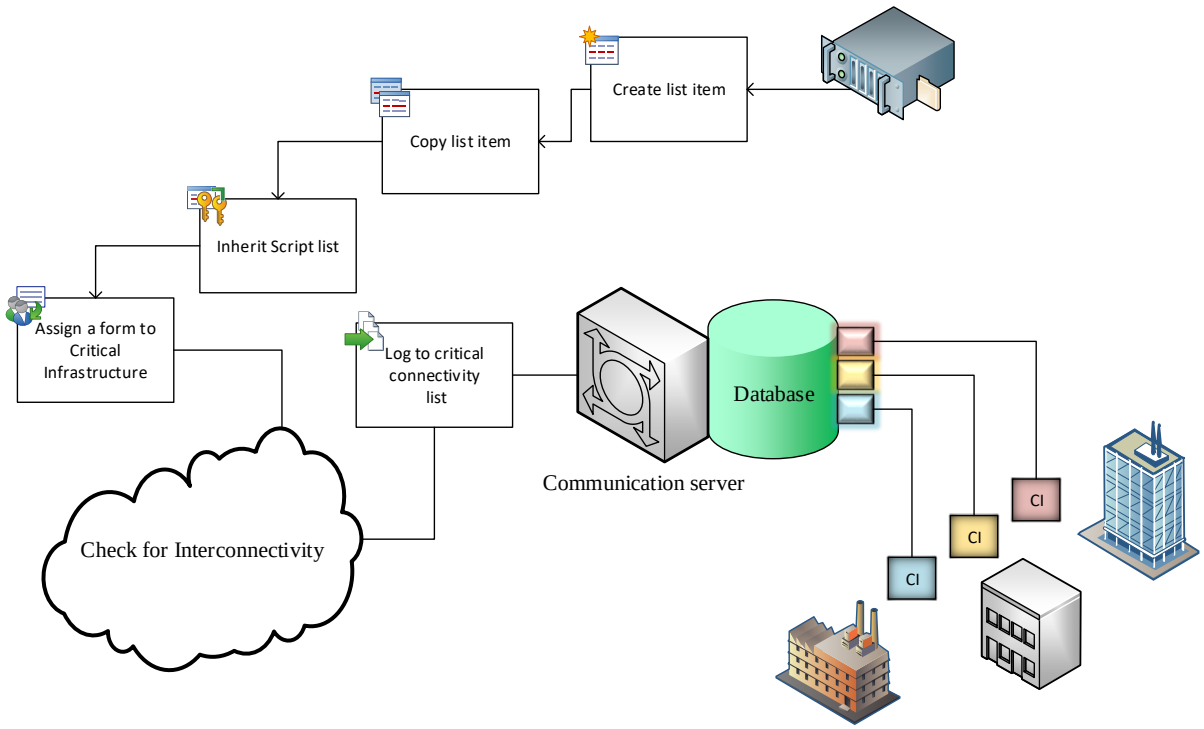


Figure 15 CIAIRS Interconnectivity Check Process

Each Critical Infrastructure needs to register its ID in a dictionary server in order to recognise all interdependency connections. Figure 16 shows the interconnectivity registration registering process.

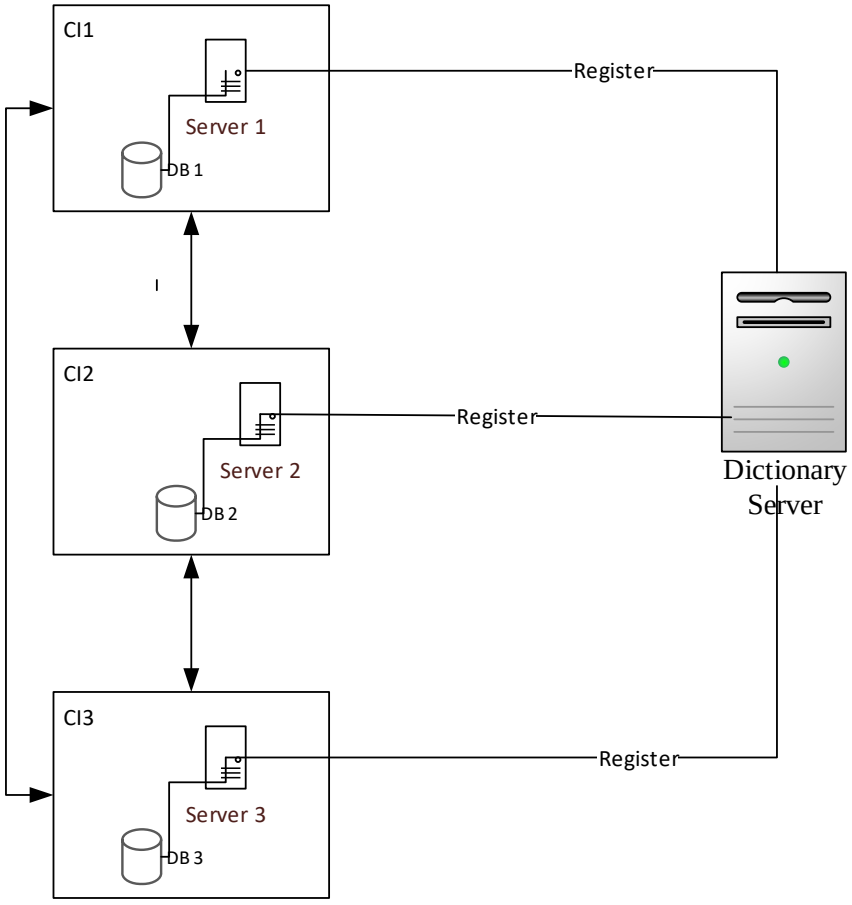


Figure 16 CIAIRS Interconnectivity Registration

4.4.7 Communication

Depending on the decision from the interconnectivity registration process, the network uses the connectivity between infrastructures to share the new abnormal behaviour with interconnected partners. This would assist other infrastructures in planning for an emerging attack or cascading impact – at all times an administrator overseas the system functionality. Figure 17 illustrates how the CIAIRS broadcasts events between the different Critical Infrastructures. Different information is sent from CIAIRS as charts, recommendations or patterns; signatures to alarm the rest of the Critical Infrastructures. In addition, CIAIRS communicates with the same plug-in system, which is presented in Figure 17. This step is dependent on the interconnectivity registration scheme. In

order to ensure a certain level of protection an operator is responsible for checking and auditing any information before sending it to other Critical Infrastructures.

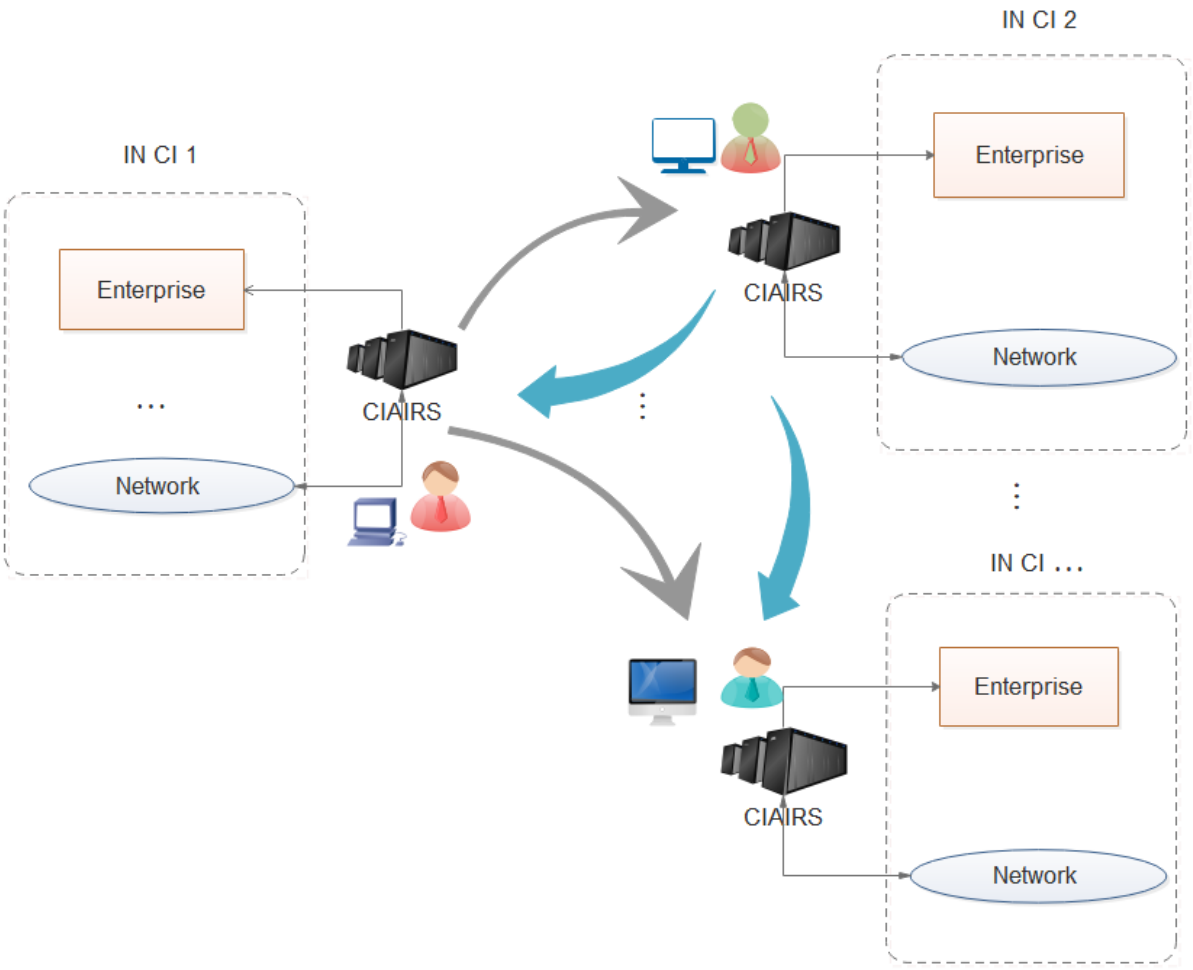


Figure 17 CIAIRS Broadcasting

Figure 18 presents the communication process for sharing attack information between the different partners. This is one of the novelties of the design. This process is divided into three main stages: decision, inner-process and action.

The first stage starts after identifying the interconnectivity between the systems. The sent script, which includes the abnormal behaviour information, ID and source are compared to the data source of each interconnected Critical Infrastructure. After that, an approval-sharing request is sent to the

Critical Infrastructure in order to continue the communication. Next, the inner-process stage starts by comparing the database source; an indicator selects the corresponding information cell and adds to the database. Based on the result, information about an anomaly is distributed, under authorization from the operator, in order to suggest the right reaction for any future attacks. Finally, an action response is sent to inform the connected Critical Infrastructure with the recommended actions to resist against the expected attack. This is known as the action stage.

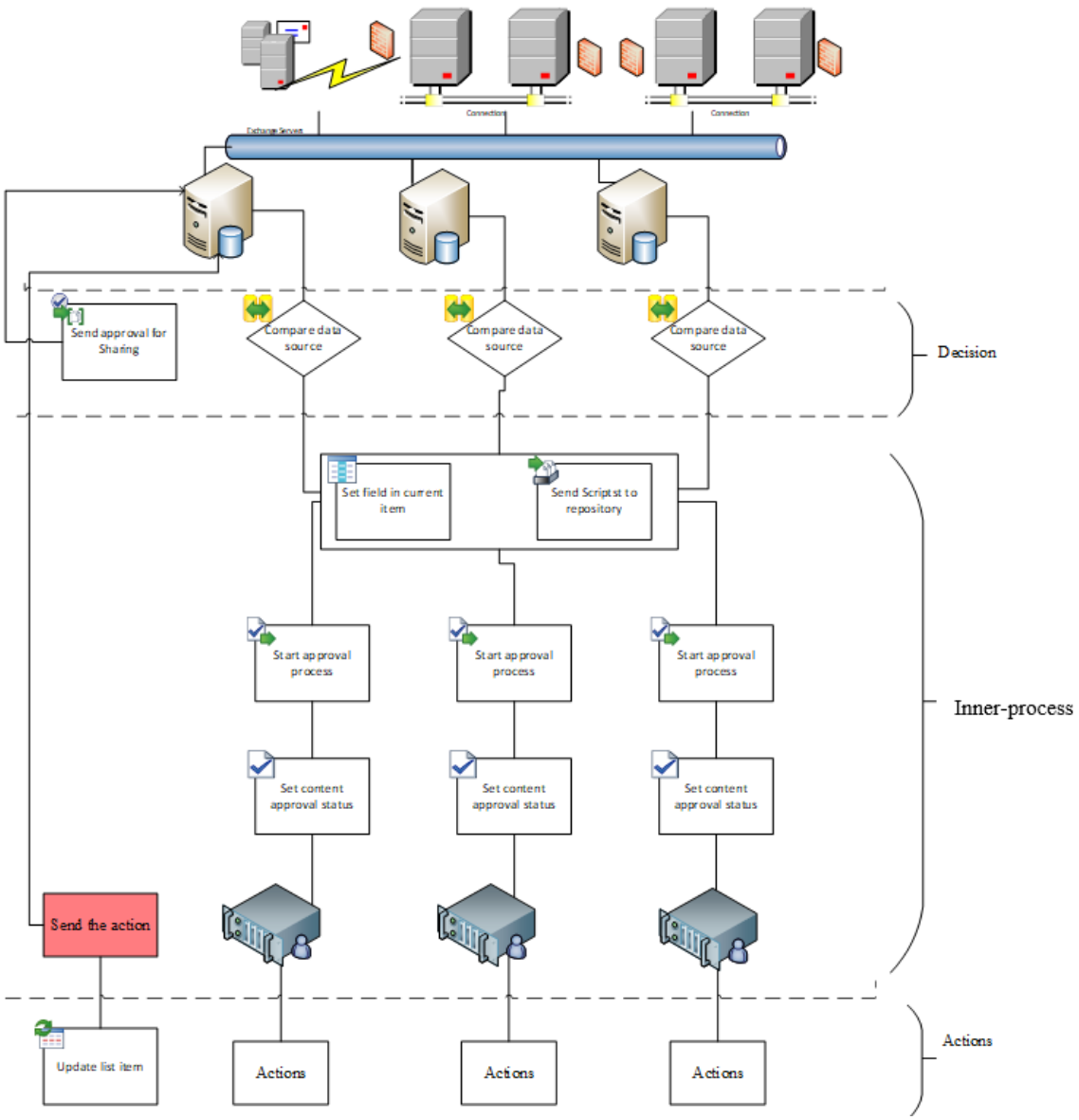


Figure 18 Communication Process

Figure 19 presents the process flow of the communication between the attacked Critical Infrastructure and the rest of the partners.

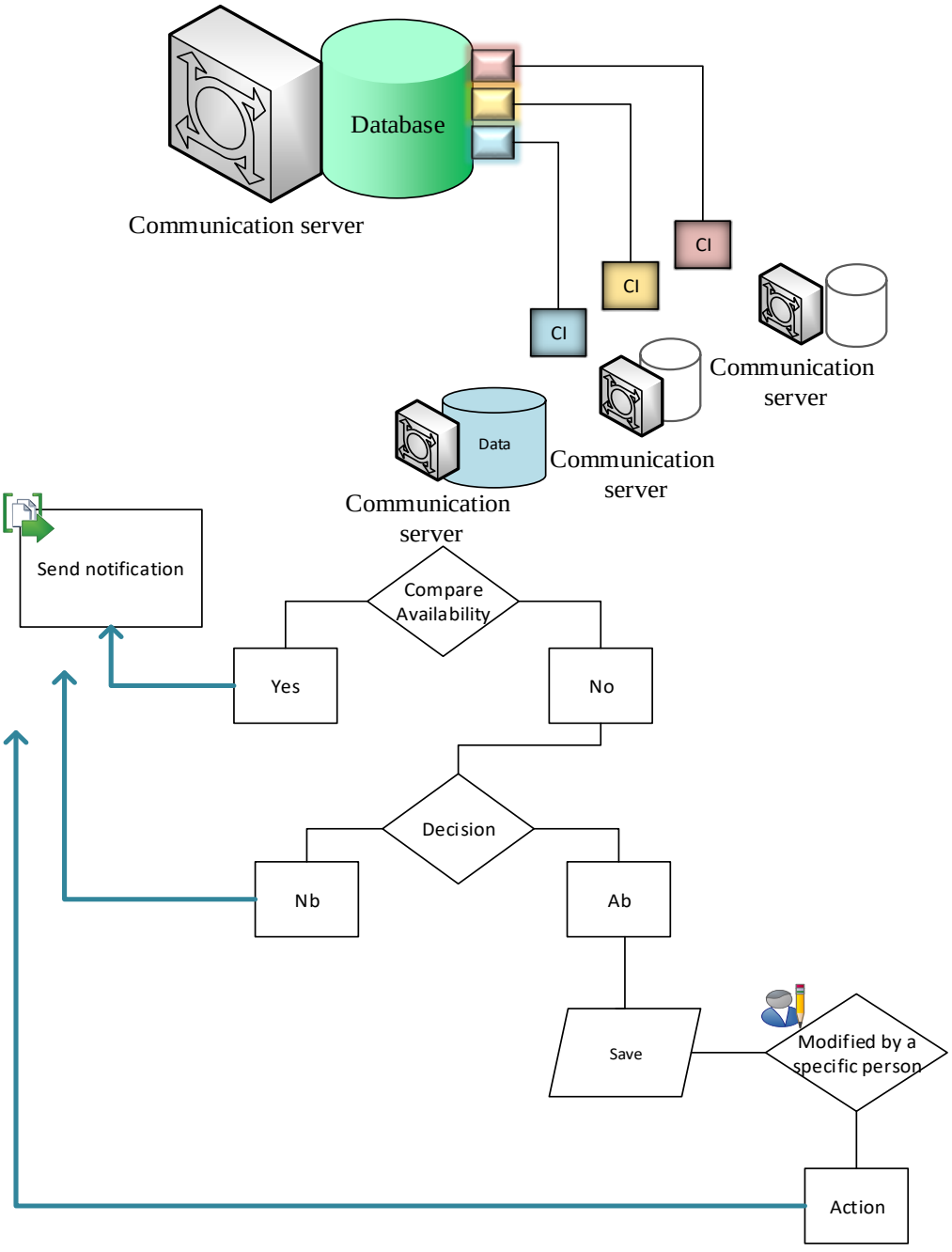


Figure 19 CIAIRS Communication Scheme

4.4.8 Action

Each Critical Infrastructure handles the attack depending on the policy that the system works under. The theoretical action scheme of handling any action is presented in Figure 20. The action towards any attack fits into two routs: the first rout focuses on updating the list, which the Critical Infrastructures have and sending it to the CIAIRS that belongs to the system. The second rout resends a notification for the main Critical Infrastructure, which has sent the attack notification in order to state the reaction toward this type of attack. However, the reaction can be similar or different.

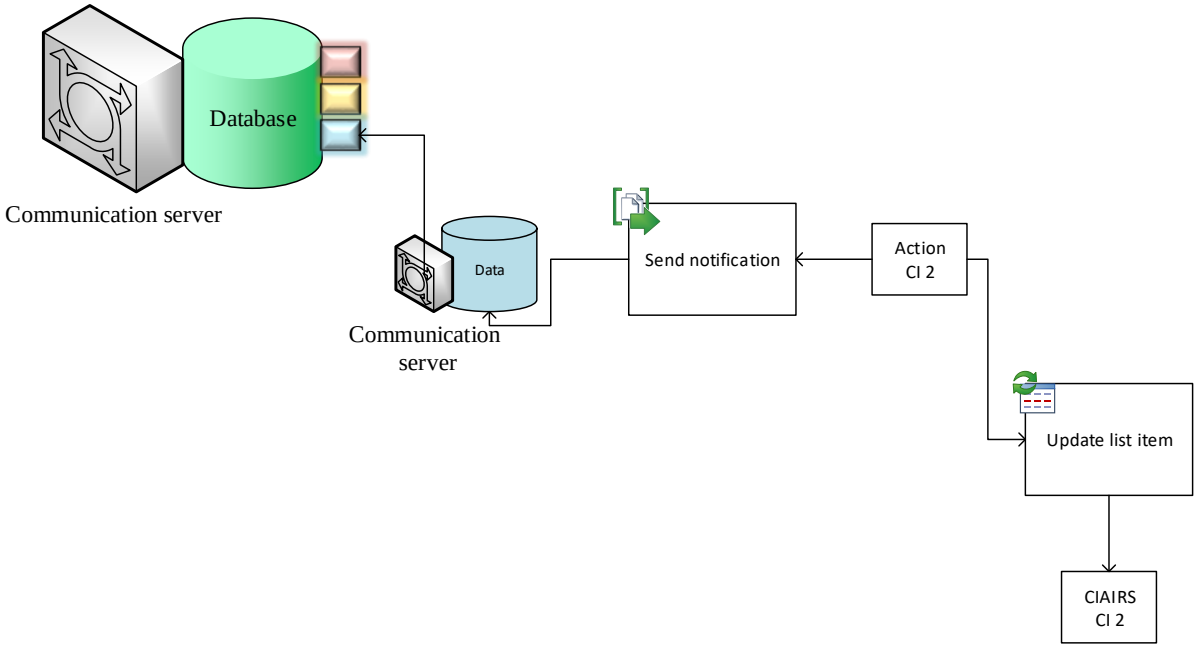


Figure 20 CIAIRS Action Scheme

4.5 Summary

This chapter started by specifying the CIAIRS framework requirements. Next, it presented a full perspective regarding the CIAIRS system, from the framework, location and communication.

Moreover, one of the main aspects in this chapter is demonstrating how the novelties of CIAIRS are able to increase the security level between the Critical Infrastructures.

CHAPTER 5

IMPLEMENTATION

5.1 Introduction

Using a semi-structured interview, a realistic example based on the Saudi Water Ministry in Jeddah was used to implement CIAIRS [135]. Realistic simulation is used to construct the data. The data is used to evaluate the system. In total, a simulation of eight Critical Infrastructures is presented for the data construction process.

5.2 Simulation Architecture

In order to evaluate CIAIRS, realistic Critical Infrastructure data is required. Therefore, eight Critical Infrastructures are constructed using Tecnomatix [41]. Implementing the simulation involved two main points: Setting up the interconnectivity at a high level, and constructing the mechanisms of each infrastructure down to a low level. Using this approach, granular datasets are constructed for the data analysis process [132]. Using this approach enables the collection of data from individual infrastructures, as individual systems, as well as from the system as a whole, which is a system of systems.

Within the simulation, failures can be introduced, which disrupt the service provision; but there is no single point of failure which can crash the simulation as a whole [136]. Therefore, this section illustrates the steps that helped in developing the eight Critical Infrastructures.

5.2.1 Map Design

The eight Critical Infrastructures presented in both Figure 21 and 22 are comprised of a Hydroelectricity plant, an Electricity Grid, a Water Distribution plant, a Sewage System, a Nuclear Power Plant, a Coal Power Plant, a Factory and a Housing complex. The selection of these Critical Infrastructures was made, as they are well-known infrastructures, present in most developed countries [137]–[139]. Two designs are chosen to present the simulation, as each of the designs display a different aspect. The perspective city map is shown in Figure 21, whereas the links between the Critical Infrastructures is illustrated by the top view scheme in Figure 22. The links are either pipes or cables.

Figure 21 Perspective City Map for the 8 Critical Infrastructures

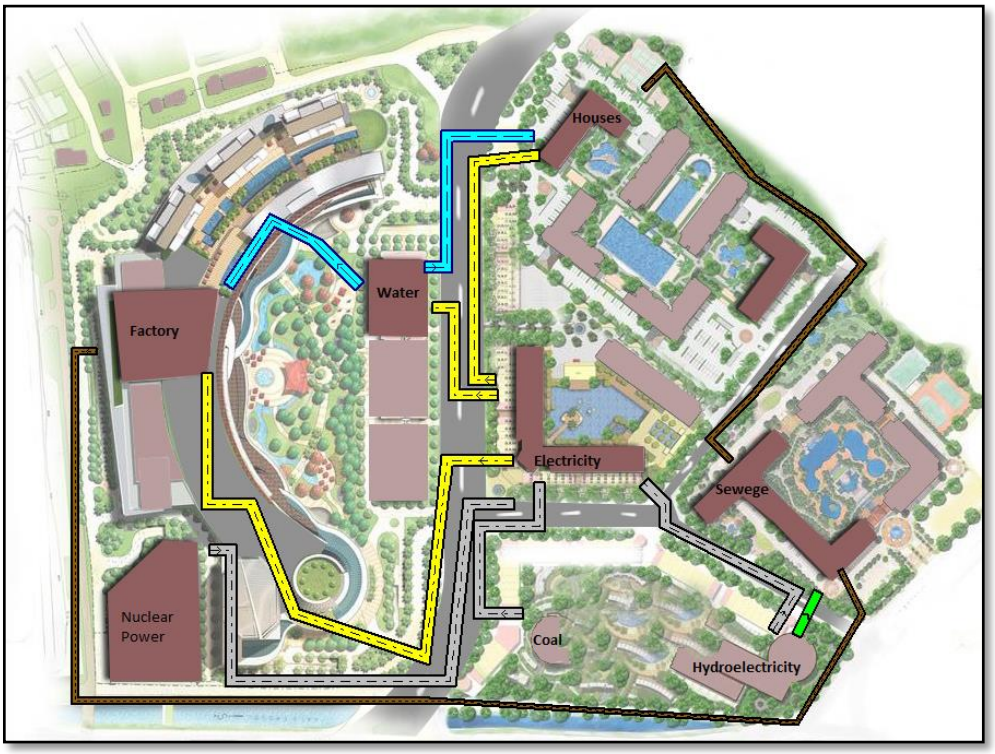


Figure 22 A Top View Scheme Illustrating the Links between the 8 Critical Infrastructures

5.2.2 Tecnomatix

The simulation is based on object-oriented modelling, where each component inserted is an individual object. Each can be adjusted and used to construct data. In order to generate simulation data, objects are created and inserted for each of the components, which together form the simulation environment. Figure 23 presents an overview of the global Critical Infrastructure system of systems, and the different supply chains, such as water pipes, electricity cables and sewage system. Each of the Critical Infrastructure systems is given a graphical icon to represent its function more clearly. They can be expanded within the simulation, to show the different objects, which comprise the system as a whole. The eight Critical Infrastructures are as follows:

1. Electricity Grid

2. Water Distribution
3. Factory
4. Nuclear Power Station
5. Coal Power System
6. Hydroelectricity system
7. Sewage System,
8. The compound of Houses.

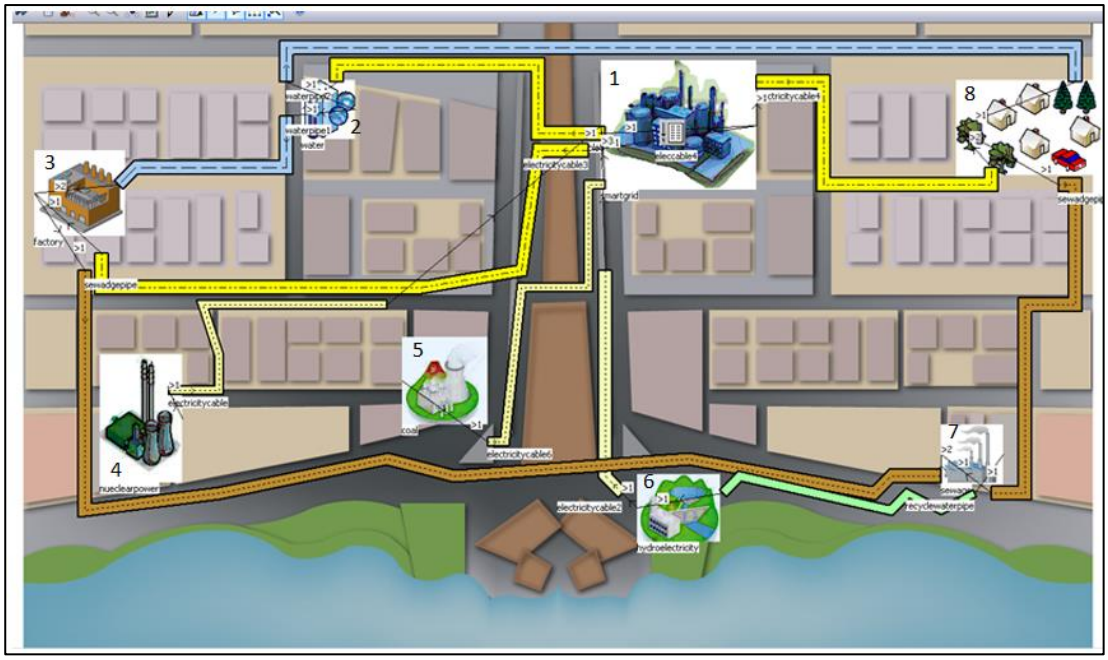


Figure 23 Simulation Plant for the Links between the eight Critical Infrastructures

The various arrangements, of the system functionality, is presented in the process flow in Figure 24.

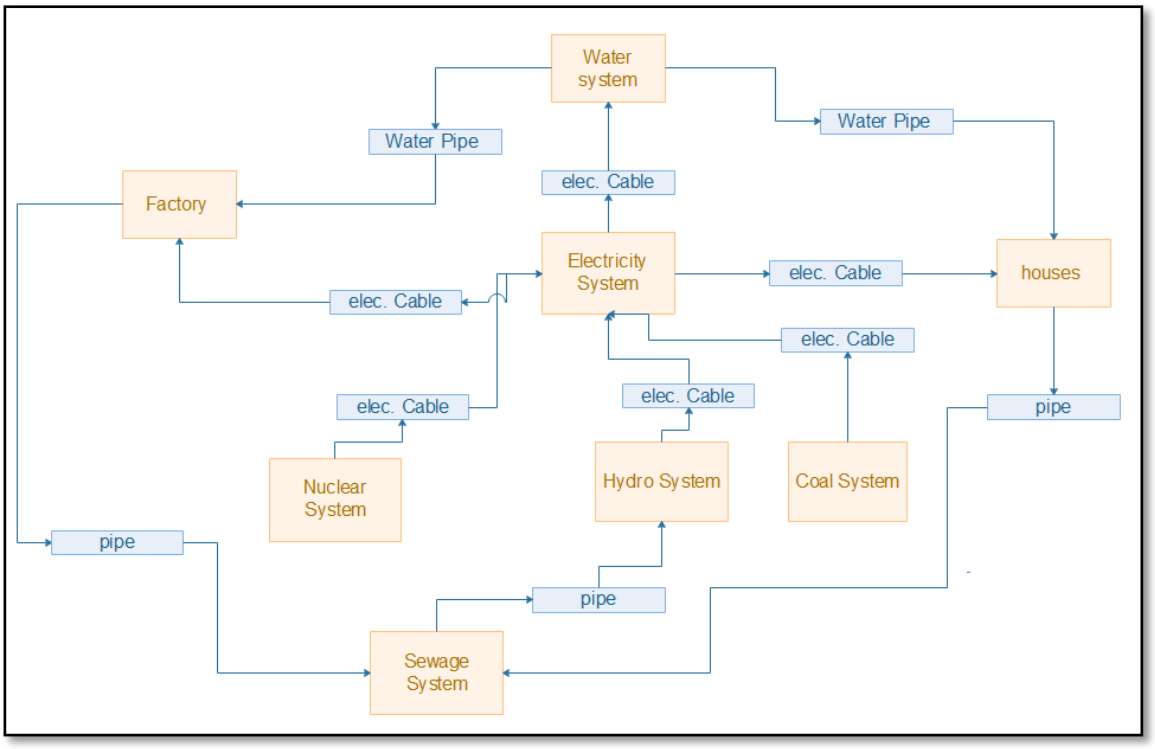


Figure 24 the Process Flow between the 8 Critical Infrastructures

5.3 Electricity System

In this subsection, the Electricity System is presented. An expanded view of the system design, the actual simulation and the process flow are depicted. Figure 25 displays the expanded view of the Electricity system. The system is linked to three sources of material, which are required for it to function. They consist of the Hydroelectricity System, Coal System and Nuclear System. In turn, the Electricity System provides electricity to three others systems; Houses, Factories and Water. Moreover, the Electricity System consists of seven different components, which are displayed in Table 5.

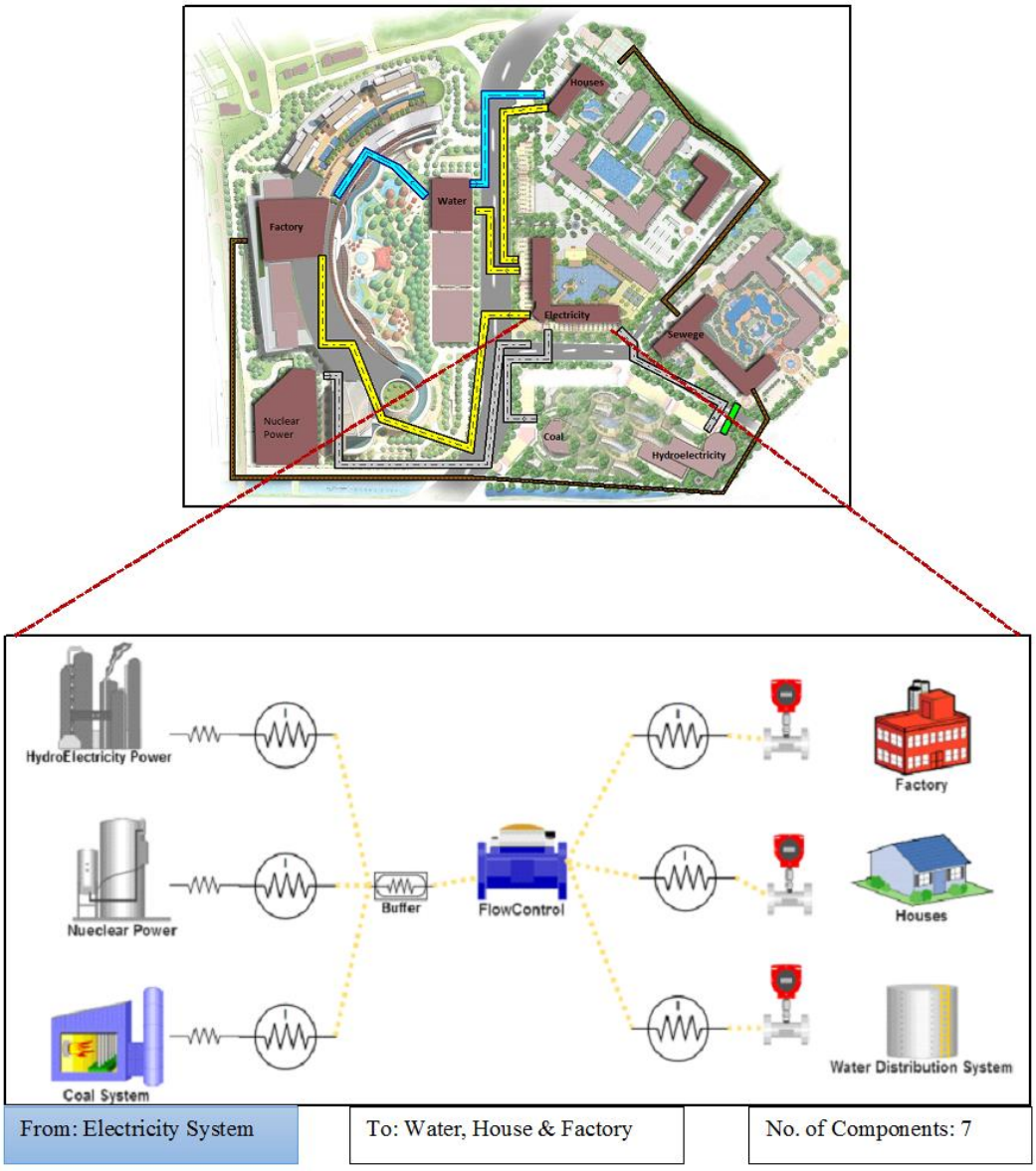


Figure 25 Electricity System Components Diagram and location within the 8 Critical Infrastructures

Table 5: Simulation Electricity Components

Components	
1. Electricity cable from the hydroelectricity to the Smart Grid in the Grid	2. Electricity cable from the Smart Grid to the Factory in the Grid
3. Electricity cable from the Nuclear to the Smart Grid in the Grid	4. Electricity cable from the Smart Grid to the Houses in the Grid

5. Electricity cable from the Coal to the Smart Grid in the Grid	6. Electricity cable from the Smart Grid to the WD in the Grid
7. the buffer in the Smart Grid	

Subsequently, Figure 26 displays the '*inside*' of the Electricity Power Plant, particularly all the components which generate electricity.

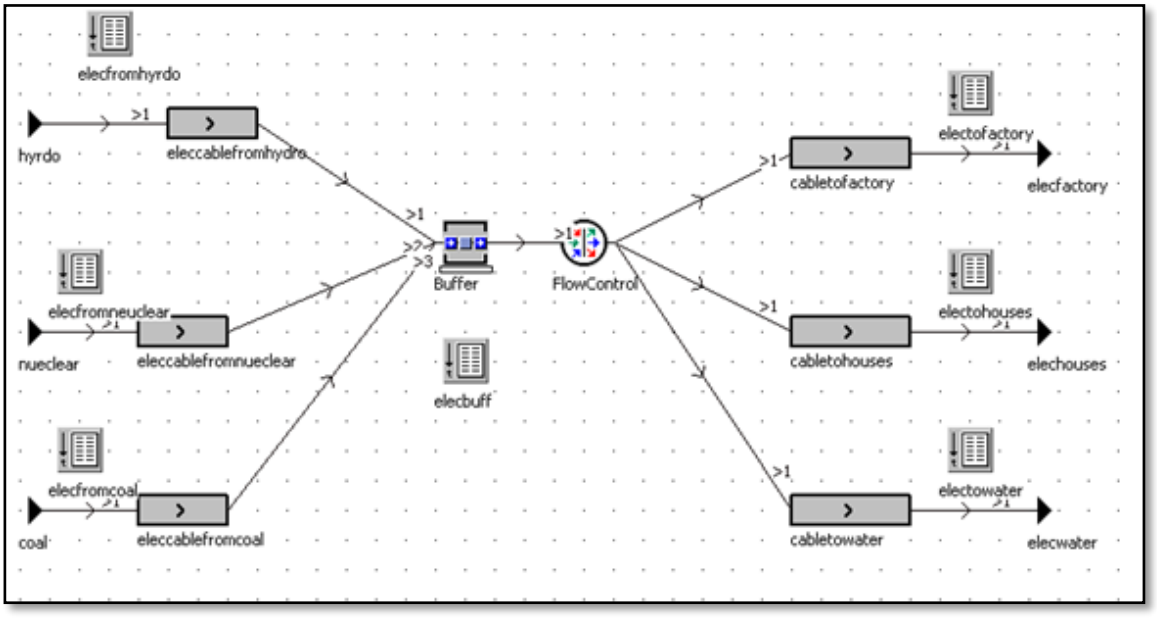


Figure 26 Simulation Plant for the Electricity System

The Electricity functionality steps are illustrated in the process flow in Figure 27.

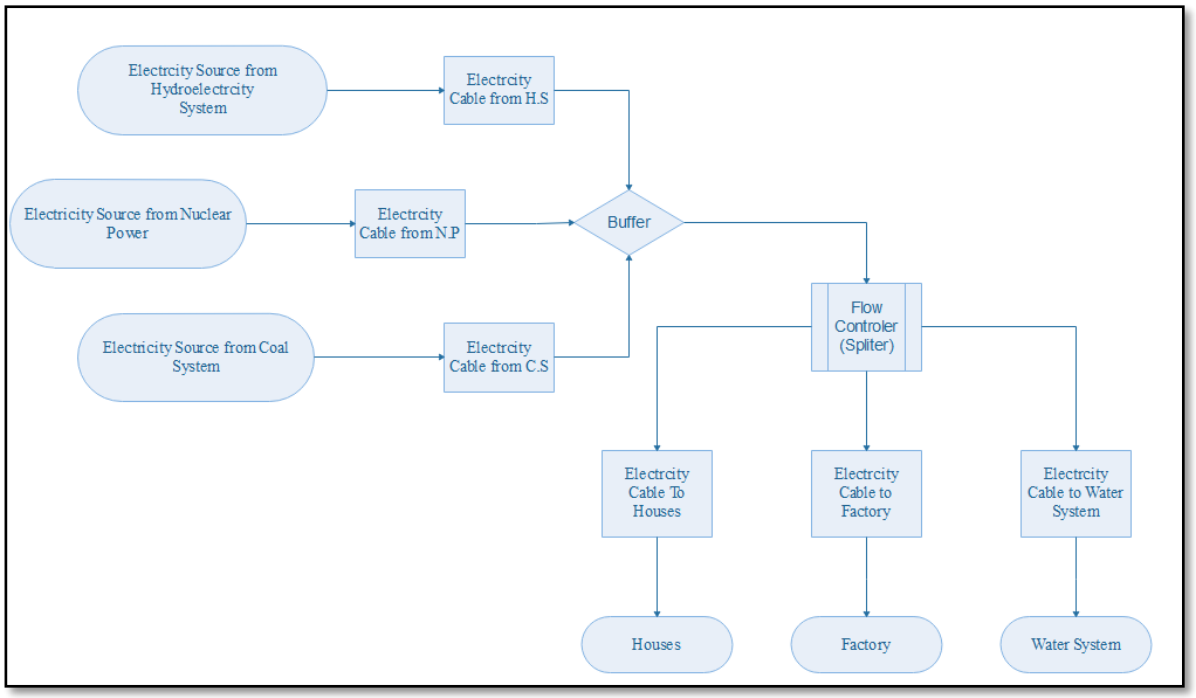


Figure 27 the Process Flow for the Electricity System

5.4 Water System

The next system is the water system, which is one of the main services that can have an impact on the housing infrastructure. Figure 28 displays the expand view of The Water Distribution System. The system links between the Electricity system and provides water to two other systems; the Houses and the Factory. Moreover, the Water System consists of 10 different components, presented in Table 6.

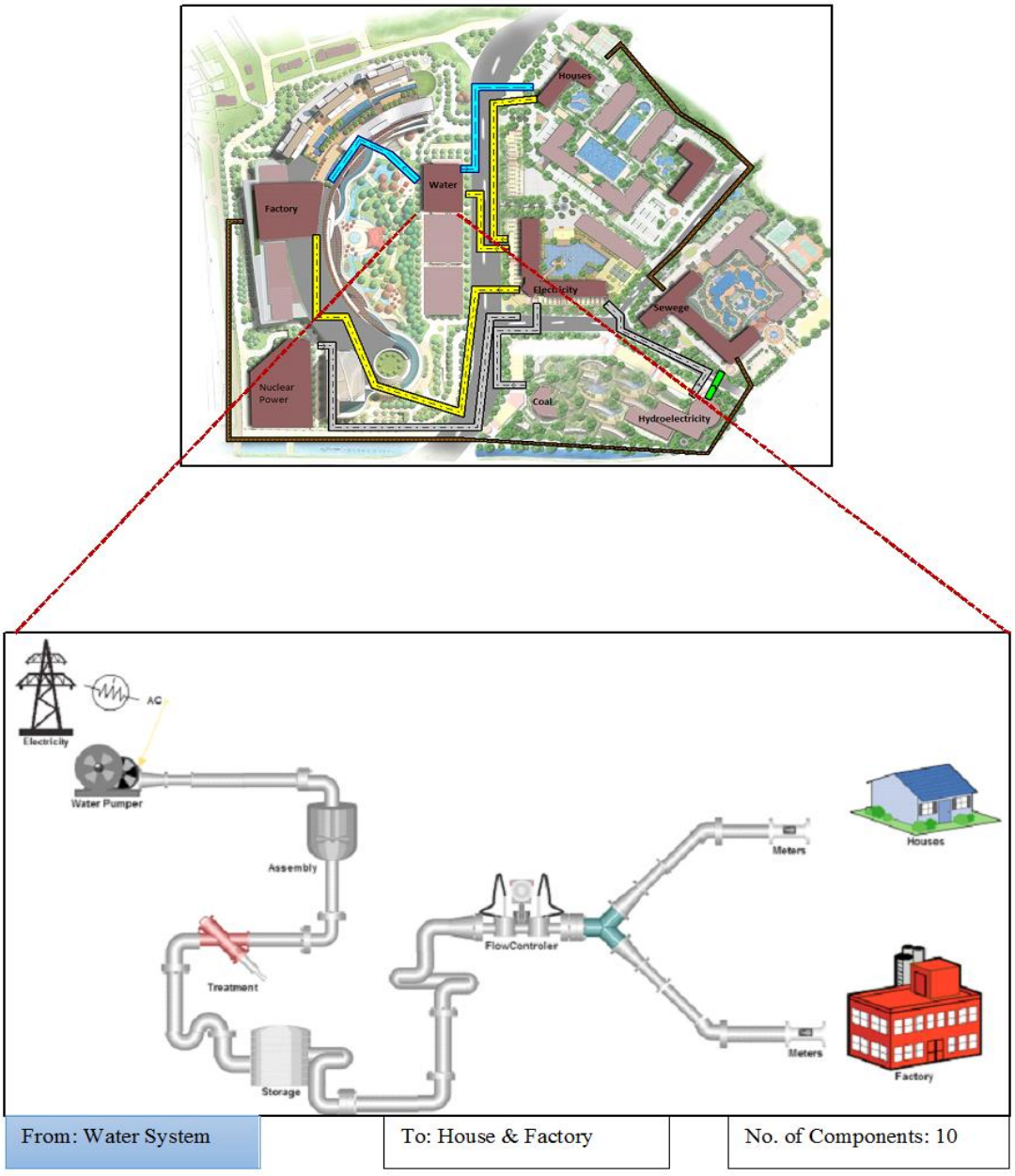


Figure 28 Water System Components Diagram and location within the 8 Critical Infrastructures

Table 6: Simulation Water Components

Components	
1. Electricity cable from the Smart Grid to the WD in the WD	2. Water pipe 3 after the treatment in the WD

3. Water pipe 1 from the source in the WD	4. WD Storage
5. WD Assembly	6. Water pipe 4 in the WD
7. Water pipe 2 in the WD	8. Water pipe from WD to Houses in the WD
9. WD treatment	10. Water pipe from WD to Factory in the WD

Figure 29 illustrates the components within the Water Distribution System. The Water Distribution System consists of a main water resource, the sea, a main electricity cable from the power plant and a transport system to send the water through pipes and feed both the houses in the compound and a factory. The Water Distribution System is controlled by a FlowControl to pump the water for both the Houses and the Factory, divided equally.

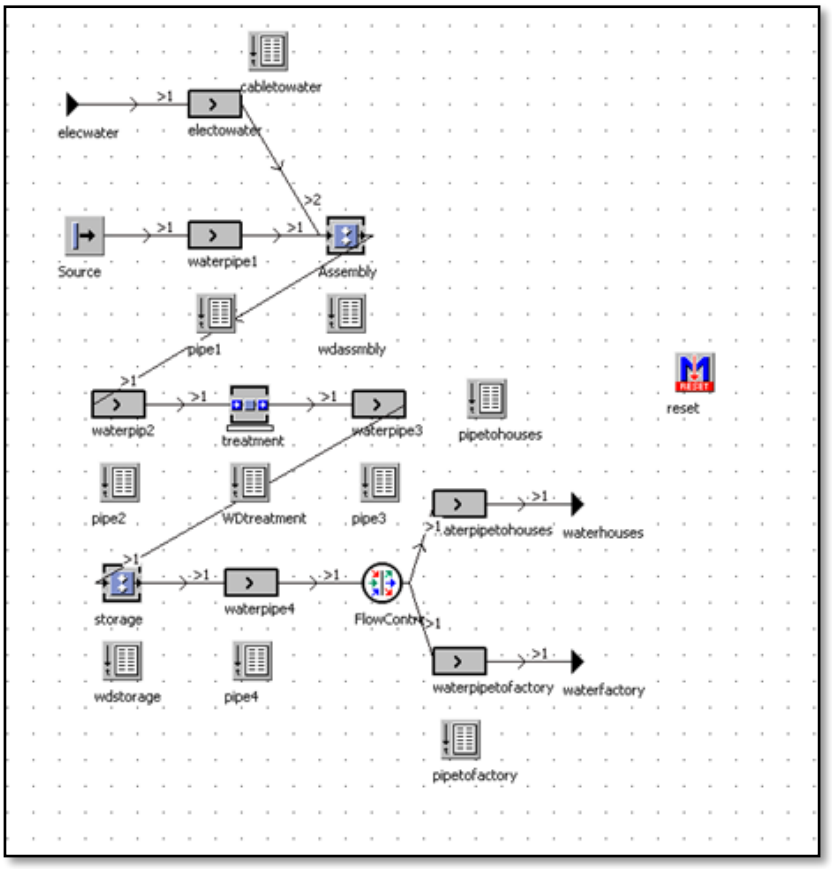


Figure 29 Simulation Plant for the Water System

Figure 30 indicates the functionality within The Water Distribution System, again through the use of a process flow diagram.

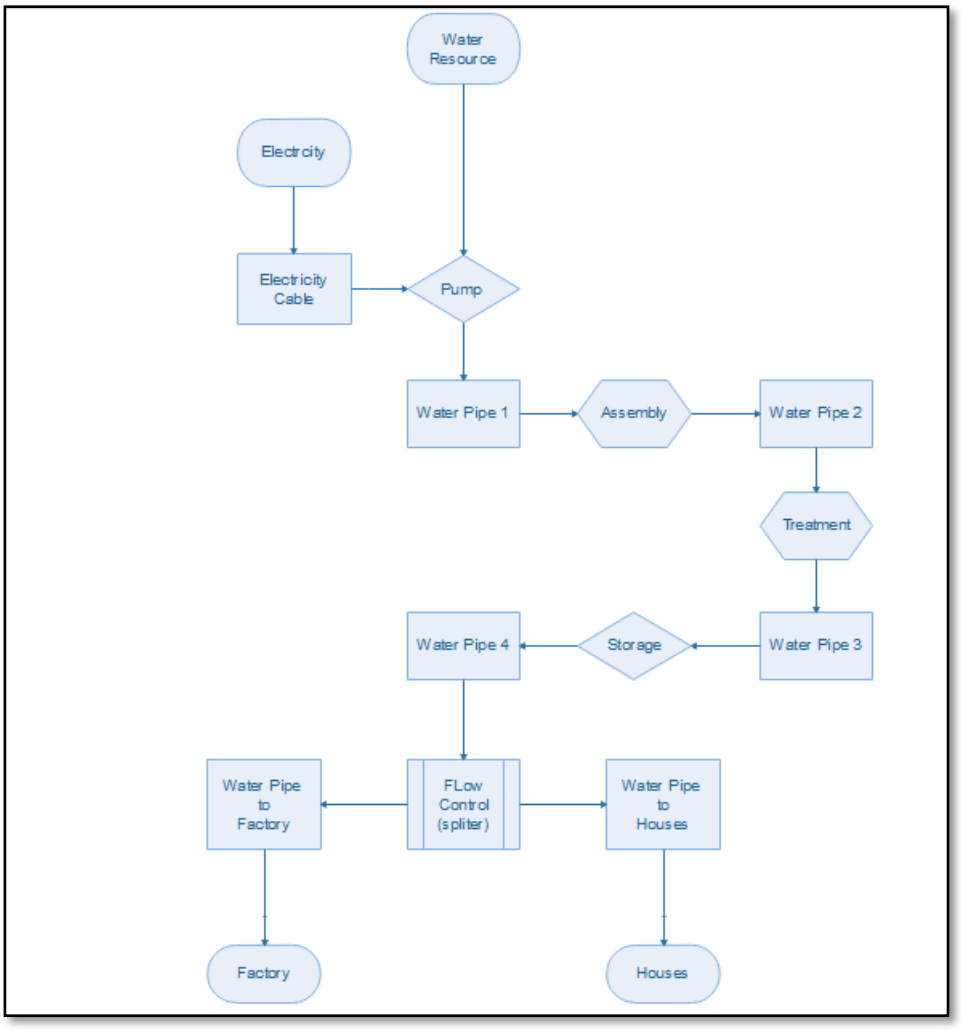


Figure 30 the Process Flow for the Water System

5.5 Factory

Next, the factory is presented. Figure 31 displays the expanded view of the three different Critical Infrastructures linked to the Factory System. Two Critical Infrastructures, the water and the electricity, supply the system so that it can provide its services. Waste from the factory is dispatched to the Sewage System. Moreover, the Factory System consists of five different components presented in Table 7.

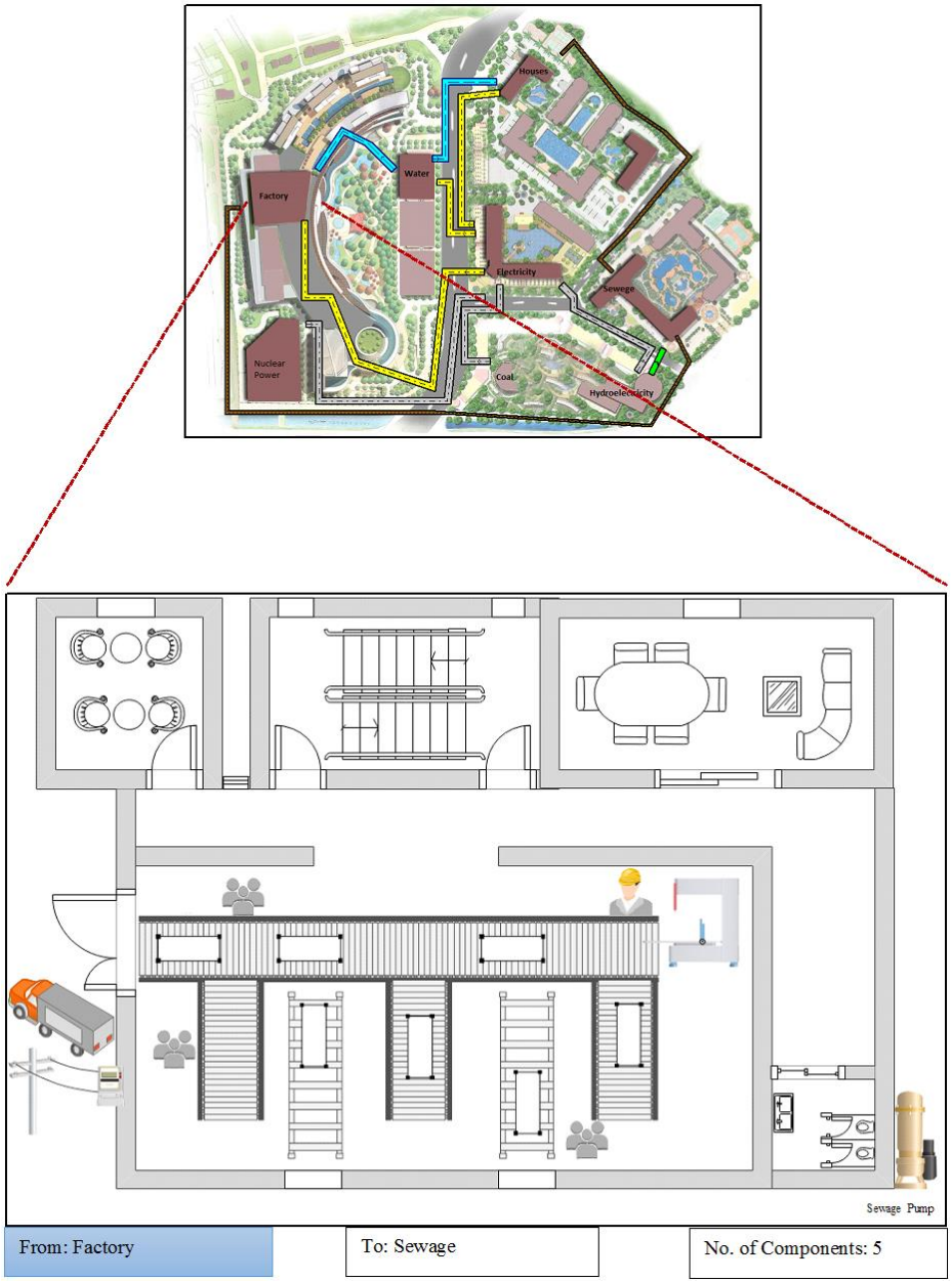


Figure 31 Factory System Components Diagram and location within the 8 Critical Infrastructures

Table 7: Simulation Factory Components

Components	
1. Single process	2. Mixture
3. Single process 1	4. Sewage pipe
5. Water pipe	

Figure 32 illustrates the components within the Factory System. The Factory System uses electricity in order to operate the machines; it also uses a water pipe from the main Water System and transfers it to the Sewage System by pipeline.

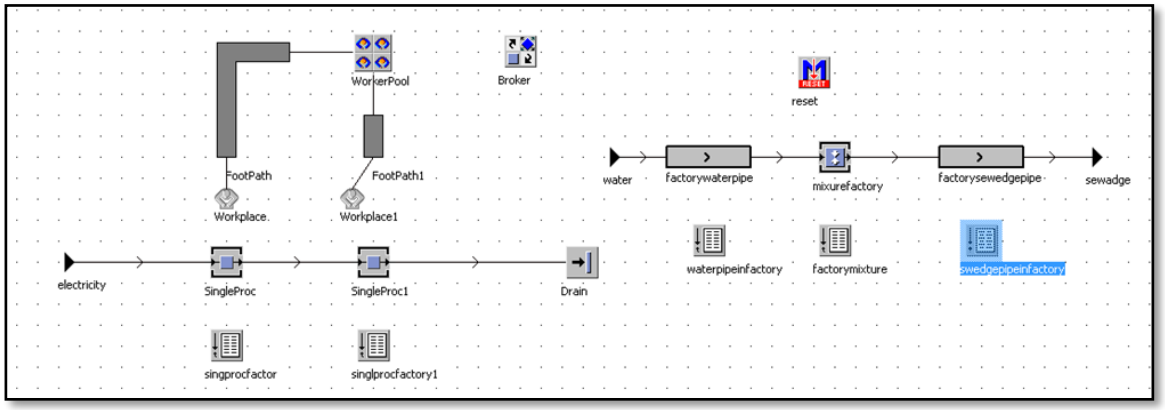


Figure 32 Simulation Plant for the System

Figure 33 indicates the functionality of the two parallel processes, (the Water System and the Electricity System) that are work within the Factory.

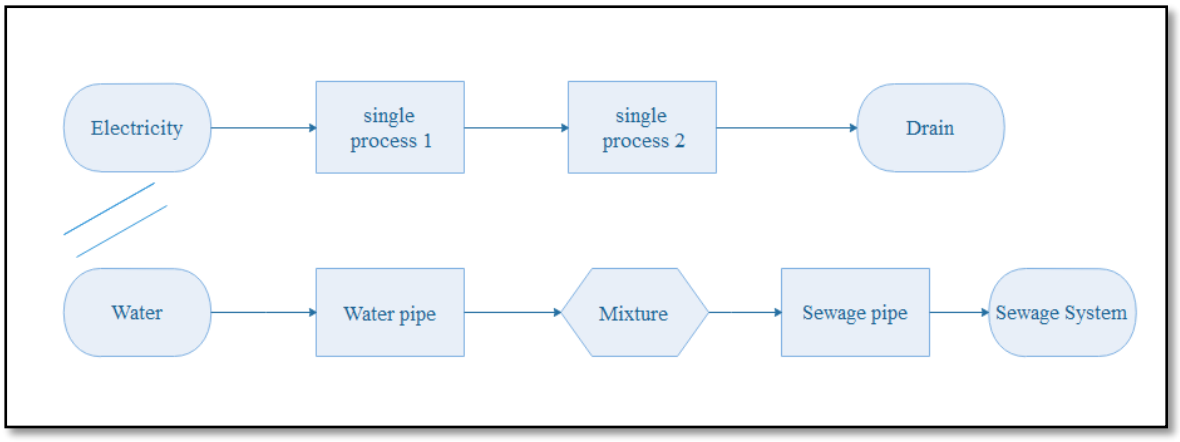


Figure 33 the Process Flow for the Factory

5.6 Nuclear Power System

In this subsection, one of the electricity sources, the Nuclear Power System, is explained. Figure 34 displays the expanded view of the Nuclear Power System. The system links only to the Electricity systems. Moreover, the Nuclear System consists of 30 different components and 4 inner systems that are indicated in Table 8.

Table 8: Simulation Nuclear Power Components

Components			
1.pump 1 in the main Nuclear Power	2. water pipe from the nuclear pump2	3. the fuel in the nuclear control rods	4. warm water pipe in the main nuclear power
5.cooling water pipe in the main nuclear power	6.water pipe in containment	7.control rods pressures	8.cooling nuclear tower

9.the steam line in the nuclear condenser	10. Steam Generator in containment	11. steam line 1 from the containment to the turbine	12. vapour pipe in the nuclear power
13. cool water pipe in the condenser	14. steam line in containment	15. steam line1 in the nuclear turbine	16. generator
17. dismantle station in the condenser	18. reactors	19. the Nuclear Turbine	20. transformer
21. cool water pipe 2 in the condenser	22. pipe1 in control rods	23. dismantle station in the nuclear turbine	24. electricity cable from the nuclear power to the Smart Grid
25. warm water pipe in the condenser	26. containment pump	27. steam line 2 in nuclear turbine	
28. pump 2 in the main nuclear power	29. pipe 2 in the control rods	30. steam line 2 from the nuclear turbine to the nuclear condenser	

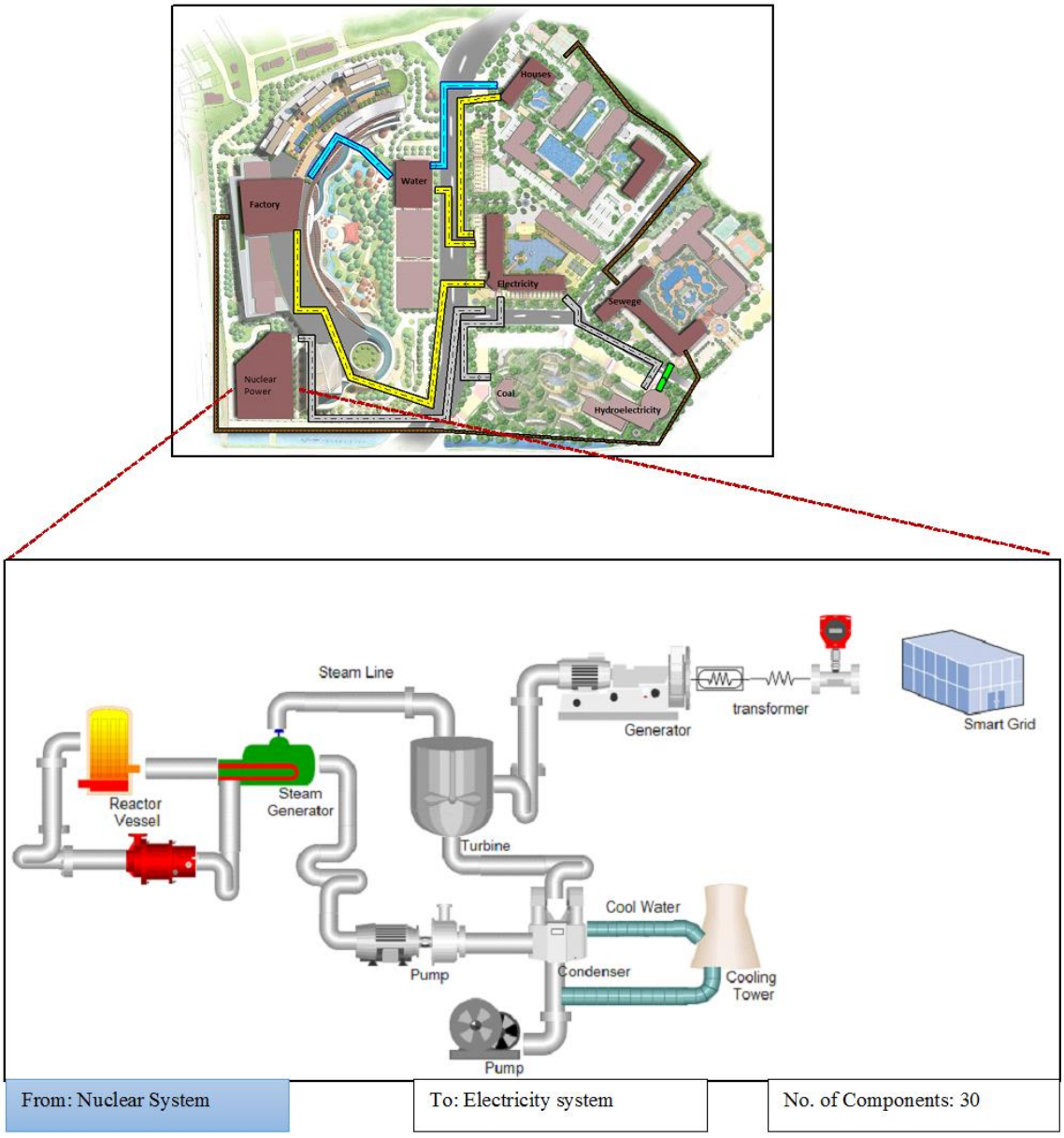


Figure 34 Nuclear System Components Diagram and location within the 8 Critical Infrastructures

Figure 35 and 36 illustrates the components within the Nuclear Power System and the Inner Nuclear System. The system consists of a pump to send cooling water to the condenser through pipes. Another pump is responsible for sending the water from the condenser through the water pipe to a steam generator. Steam lines are used to send the steam between the reactor and the

turbine. Finally, the nuclear system turbine sends the steam to the generator in order to transfer energy to the electricity grid infrastructure.

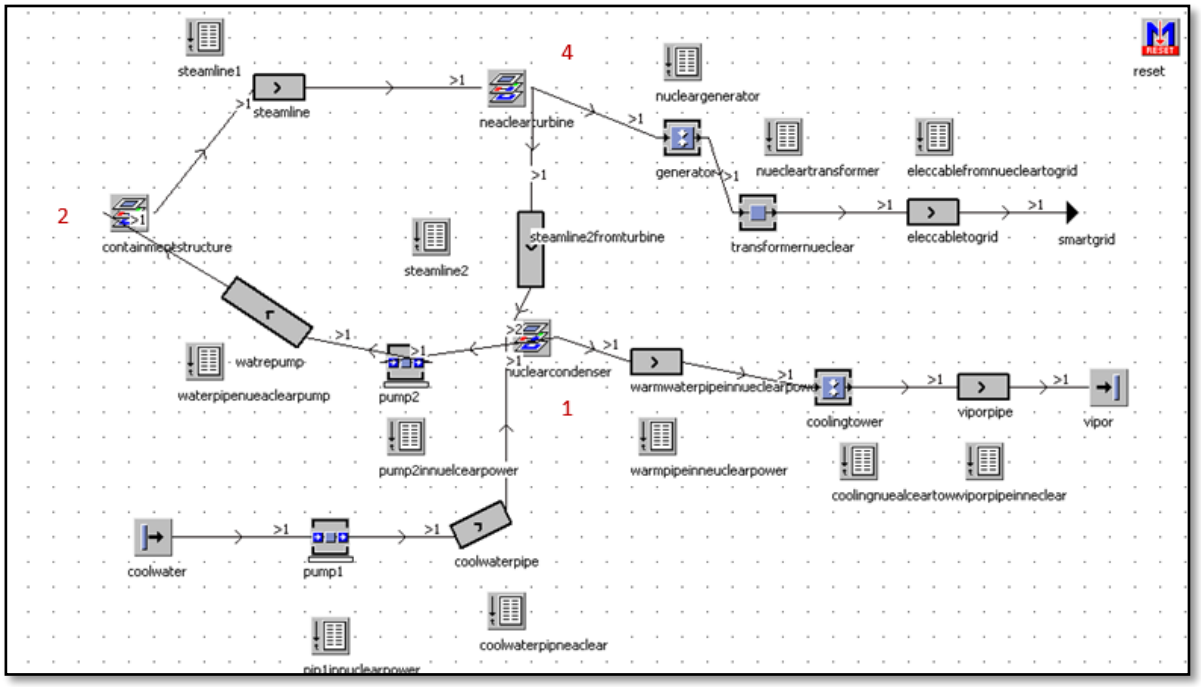


Figure 35 Simulation Plant for the Nuclear System

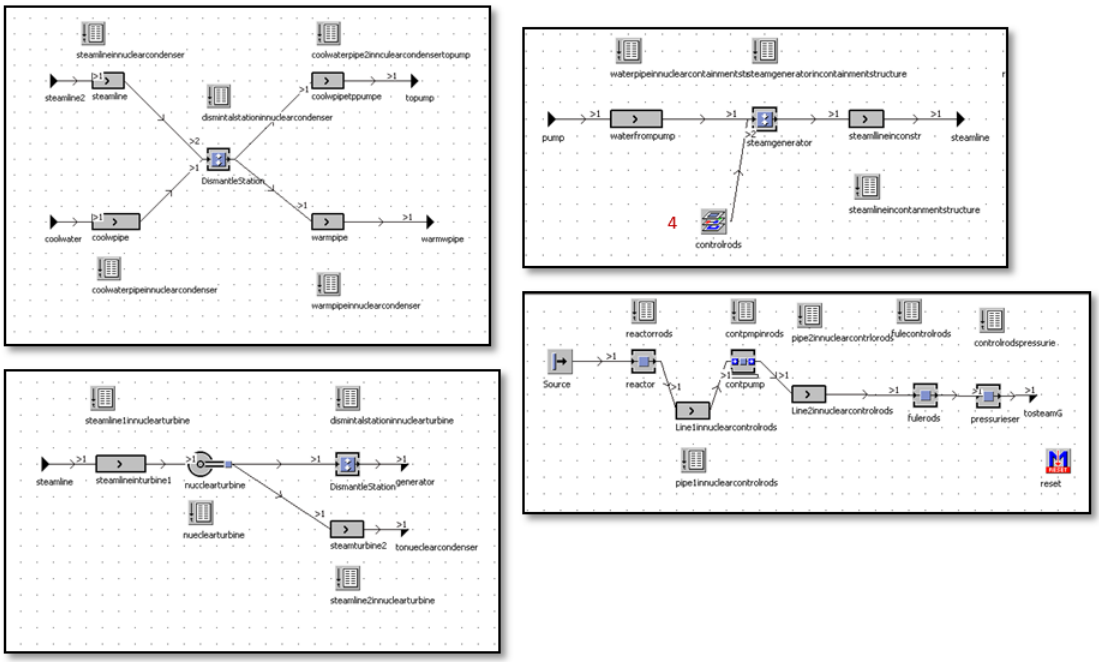


Figure 36 Simulation Plant for the Nuclear Inner System

The Nuclear Power System functionality is illustrated in more detail, by the process flow in Figure 37.

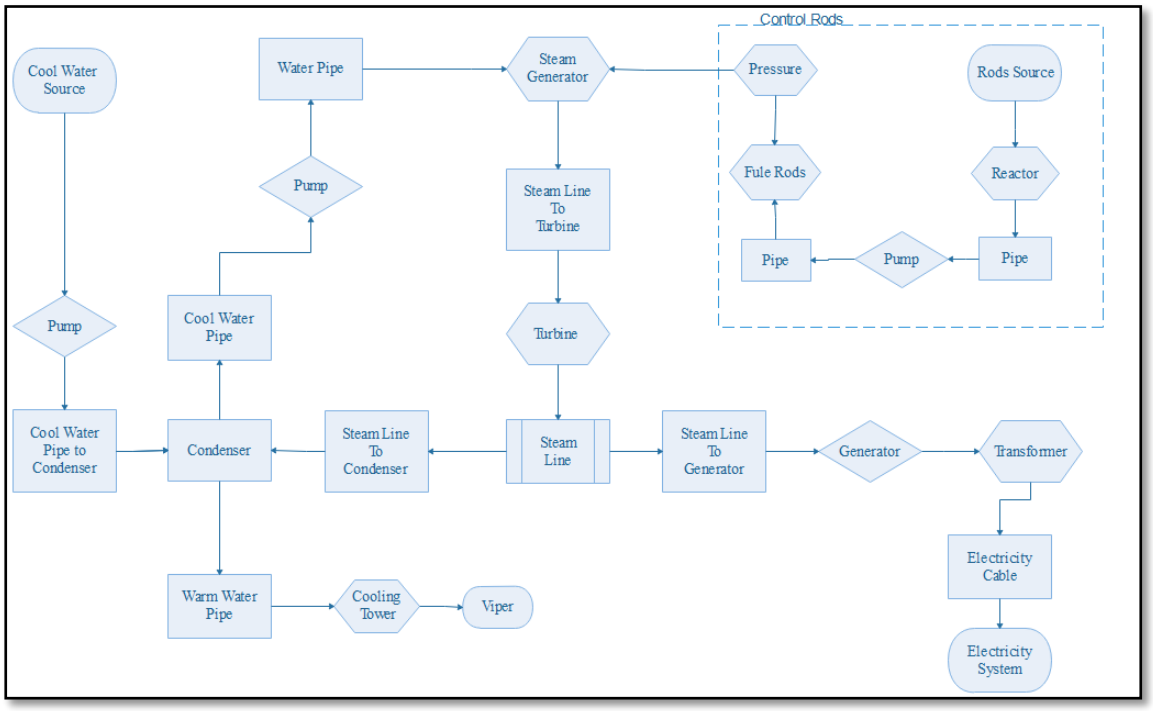


Figure 37 the Process Flow for the Nuclear System

5.7 Coal System

The second electricity source is the Coal System. Figure 38 displays the expanded view of the Coal System. The system only links to the Electricity systems. Moreover, the Coal System consists of 22 different components and 3 inner systems, which are presented in Table 9.

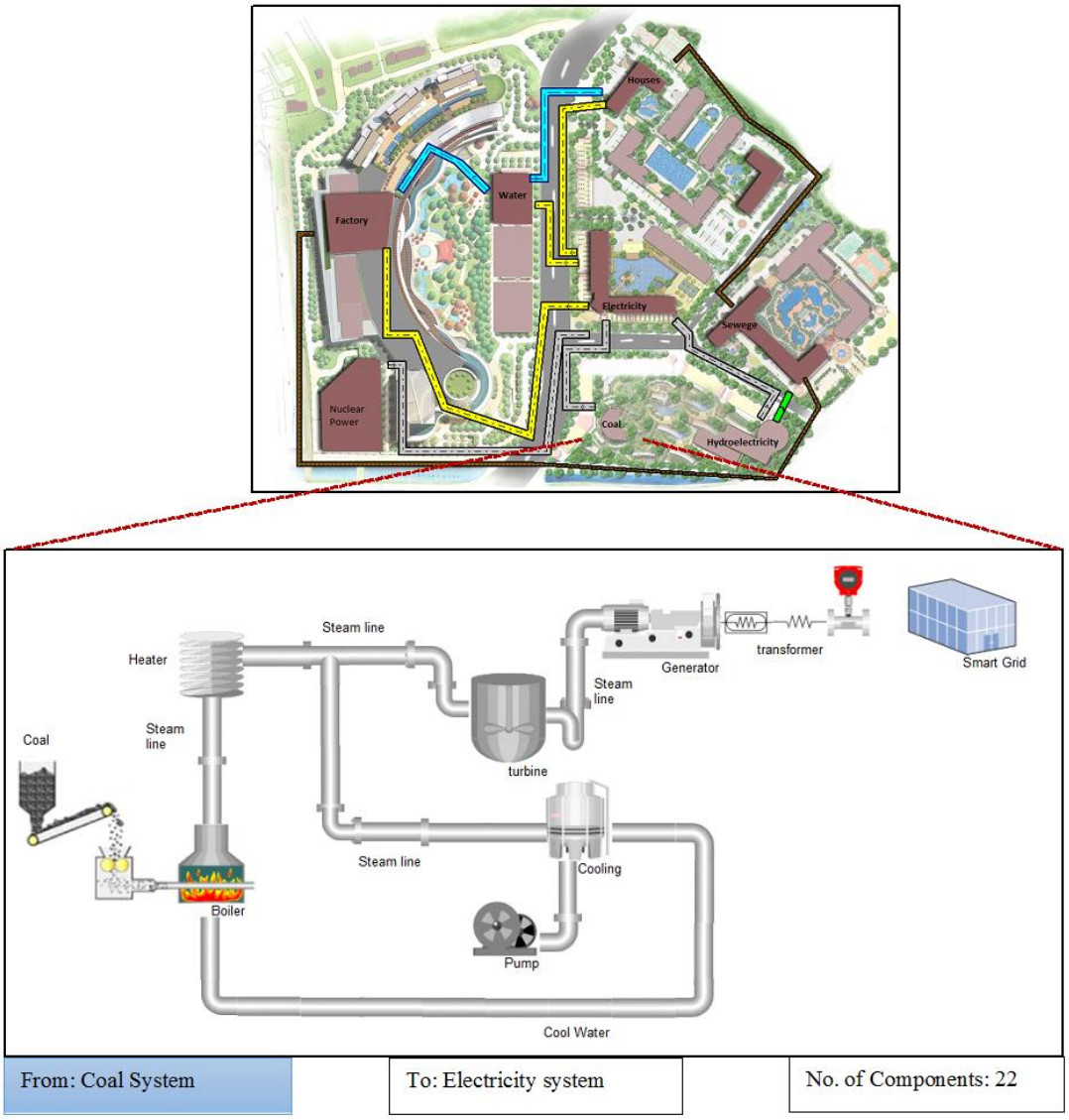


Figure 38 Coal System Components Diagram and location within the 8th Critical Infrastructures

Table 9: Simulation Coal Components

Components		
1. coal tunnel to coal boiler	2. the coal turbine	3. the condenser cooling in the coal

4. coal tube in the boiler	5.a steam pipe in turbine to the condenser	6. cool water pipe in the coal condenser
7. cool water pipe in the boiler	8. dismantle station in turbine	9. cool water pipe from the coal condenser to the coal boiler
10. the coal heat tube in boiler	11. steam pipe in turbine to the generator	12. generator
13. coal furnace in the boiler	14. steam pipe from turbine to condenser	15. transformer
16. the steam pipe in the boiler	17. steam pipe from the coal turbine in the coal condenser	18. electricity cable in the coal to the SG
19. steam pipe from the boiler to the coal turbine	20. water pumper in coal condenser	
21. steam pipe in coal turbine	22. water pipe in coal condenser	

Figure 39 and 40 illustrate the components within the Coal Distrusted System and the Inner Coal System. The system consists of a boiler, heater, and a steam line to send the steam for both turbine and cooling system, and a water pump to send the water through the pipe to the cooling system. Similarly to the Nuclear Plant, the Coal System turbine sends the steam to the generator in order to generate electricity.

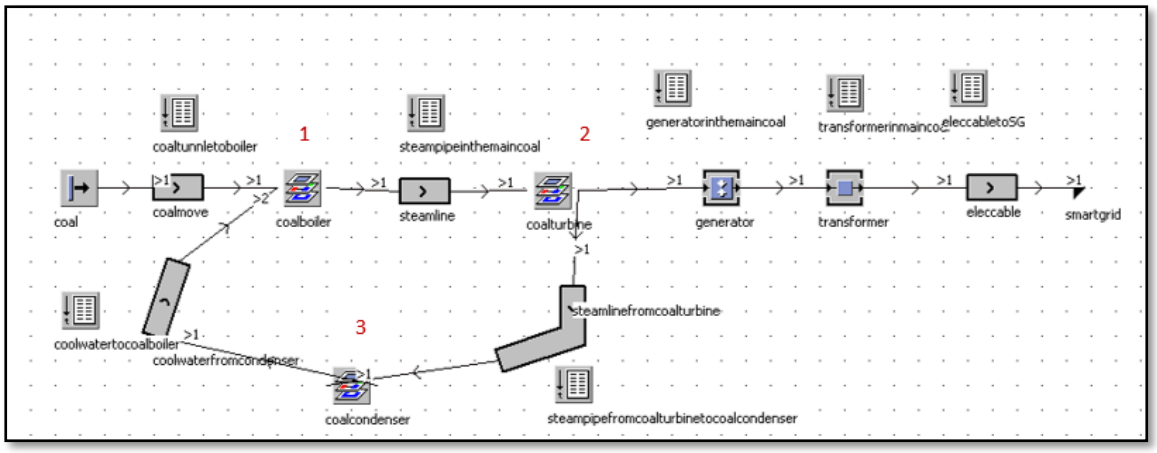


Figure 39 Simulation Plant for the Coal System

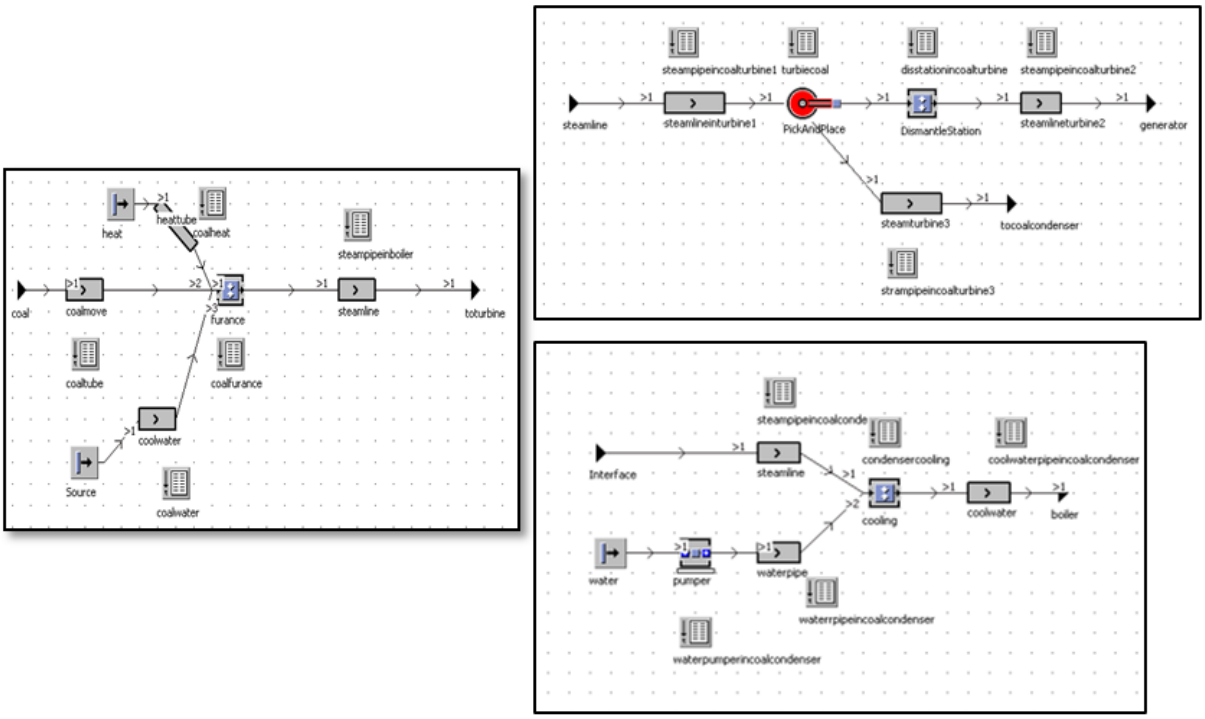


Figure 40 Simulation Plant for the Coal Inner System

The Coal process flow functionality is demonstrated in Figure 41.

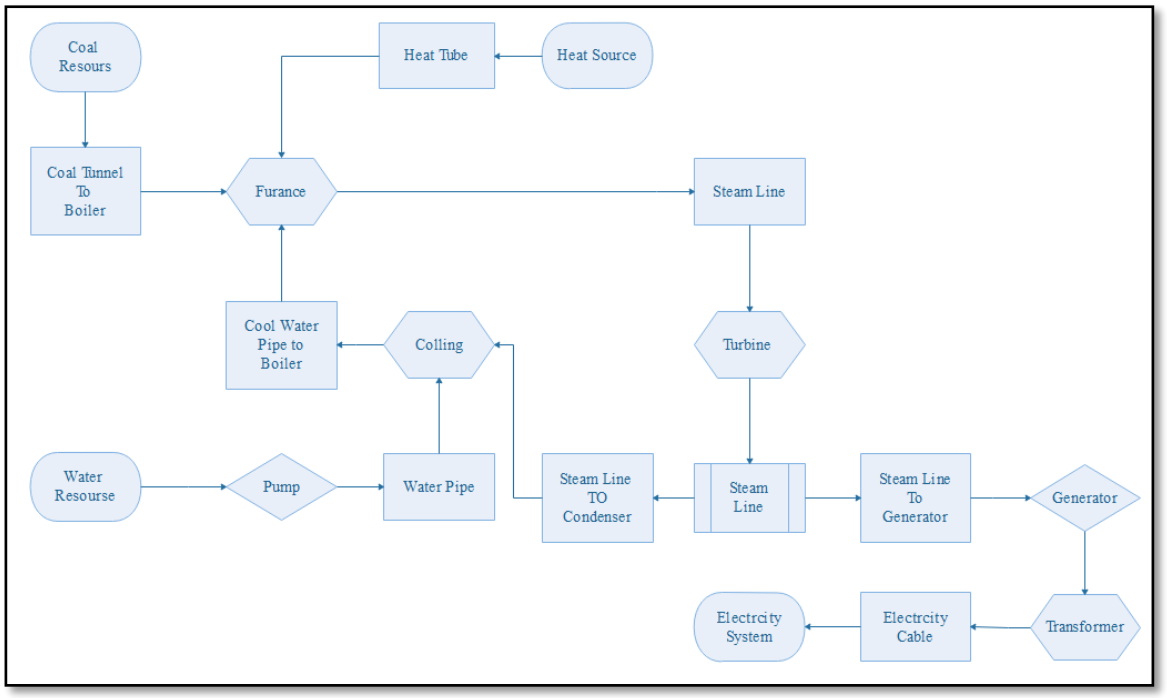


Figure 41 the Process Flow for the Coal System

5.8 Hydroelectricity System

Another electricity source is the Hydroelectricity system, which is presented in this subsection. Figure 42 displays the expanded view of the discussed Critical Infrastructure. The system links only to the Electricity systems. Moreover, the Hydroelectricity System consists of nine different components that are displayed in Table 10.

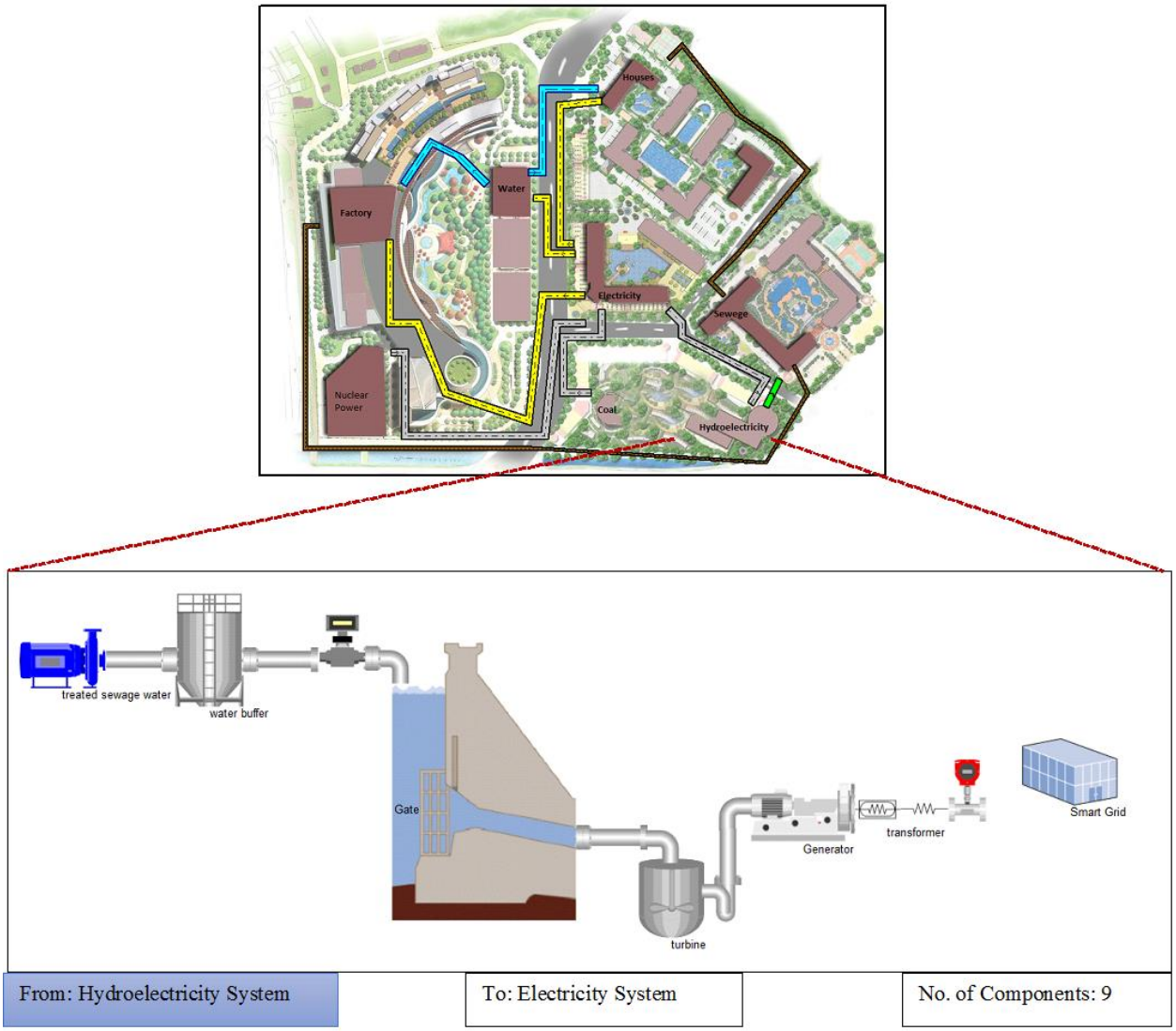


Figure 42 Hydroelectricity System Components Diagram and location within the 8 Critical Infrastructures

Table 10: Simulation Hydroelectricity Components

Components	
1. Water pipe from the sewage to the Hydroelectricity	2. the turbine
3. buffer in the Hydroelectricity system	4. generator

5. water pipe 1 from sewage in the Hydroelectricity system	6. transformer
7. sluice-gate	8. electricity cable
9. water pipe 2	

The Hydroelectricity System components consist of a water source, which is the sea, and a second source from the sewage system recycling. A sluice gate controls the water level using water pipes. These components presented in Figure 43.

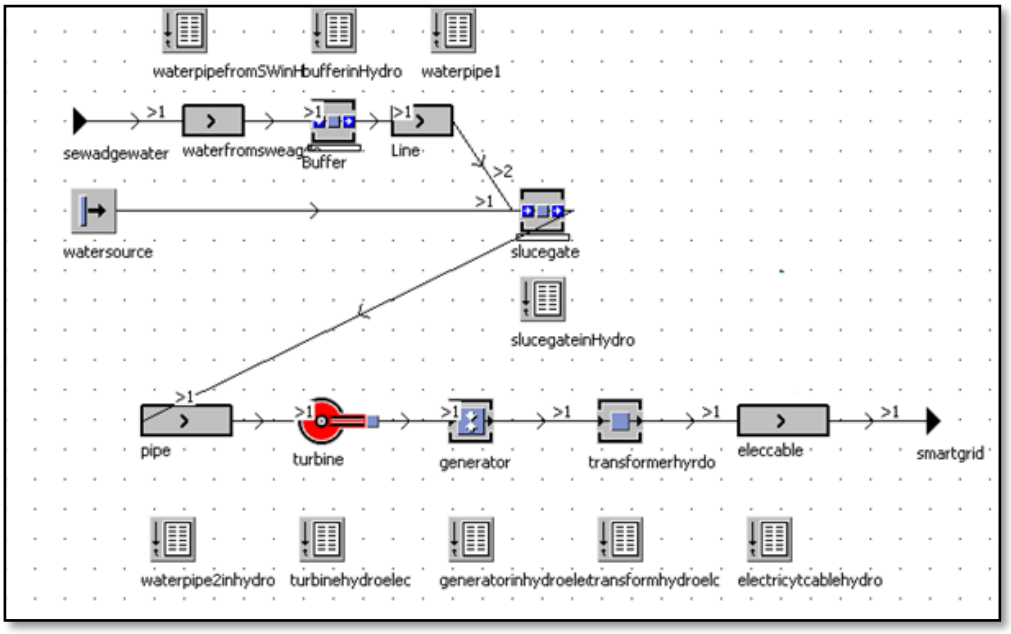


Figure 43 Simulation Plant for the Hydroelectricity System

Figure 44 indicates the Hydroelectricity process flow performance in depth.

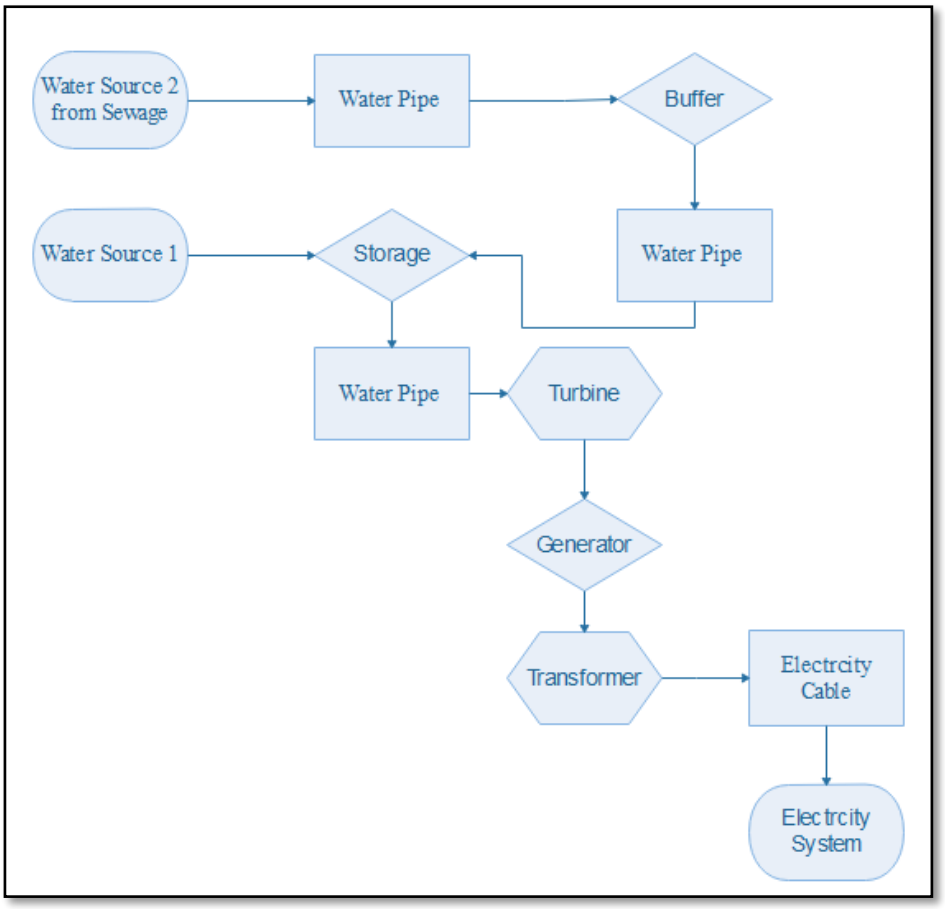


Figure 44 the Process Flow for the Hydroelectricity System

5.9 Sewage System

The final material generating infrastructure is the Sewage, and Figure 45 displays the expanded view. Three different Critical Infrastructures are linked with the Sewage System; including the Houses, the Factory and the Hydroelectricity System. The system collects the waste from both Houses and Factory System and sends re-filtered water to the Hydroelectricity System. Moreover, the Sewage System consists of eight different components that are indicated in Table 11.

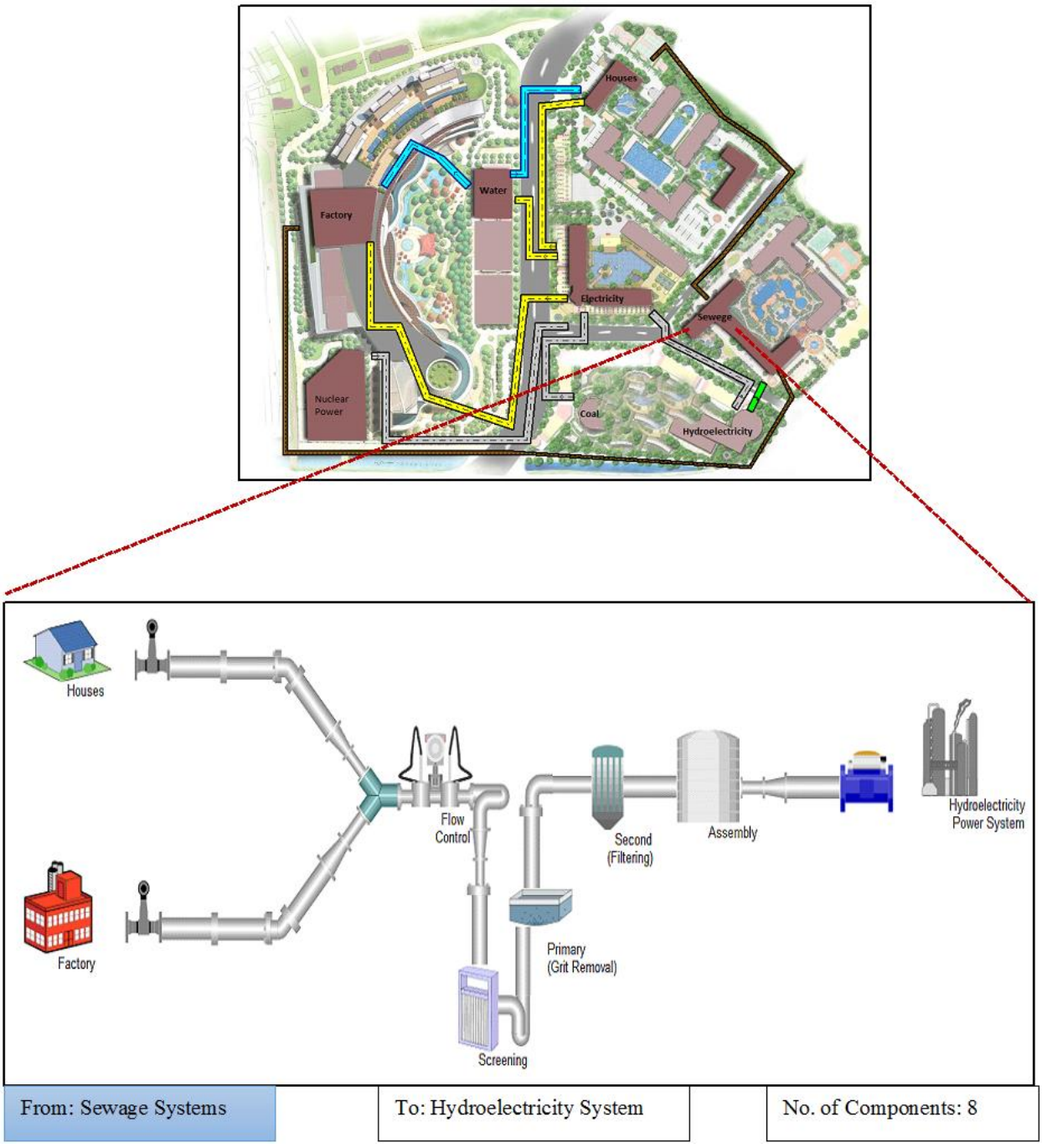


Figure 45 Sewage System Components Diagram and location within the 8 Critical Infrastructures

Table 11: Simulation Sewage Components

Components	
1. sewage pipe from houses	2. Primary process

3. sewage pipe from factory	4. second process
5. sewage pipe 1	6. assembly process
7. Screening	8. water pipe

The Sewage System components are indicated in Figure 46. A FlowController is used as a controller for sending the waste through the main pipe. The waste goes through three main processes in order to re-filter the water; these include screening, primary filtration and secondary filtration. Finally, the re-filtered water is sent to the Hydroelectricity System by a pipe.

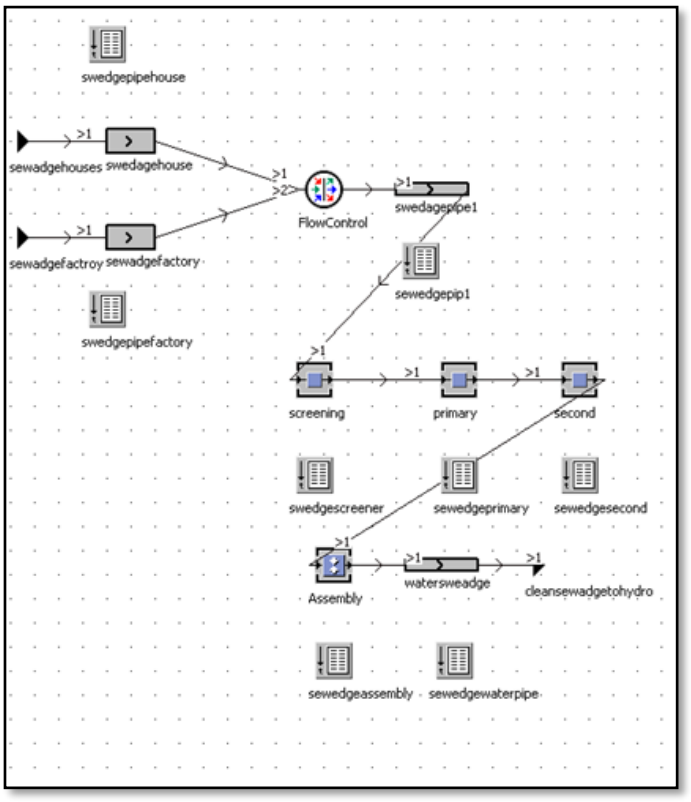


Figure 46 Simulation Plant for the Sewage System

The Sewage process functionality steps are indicated in detail in Figure 47.

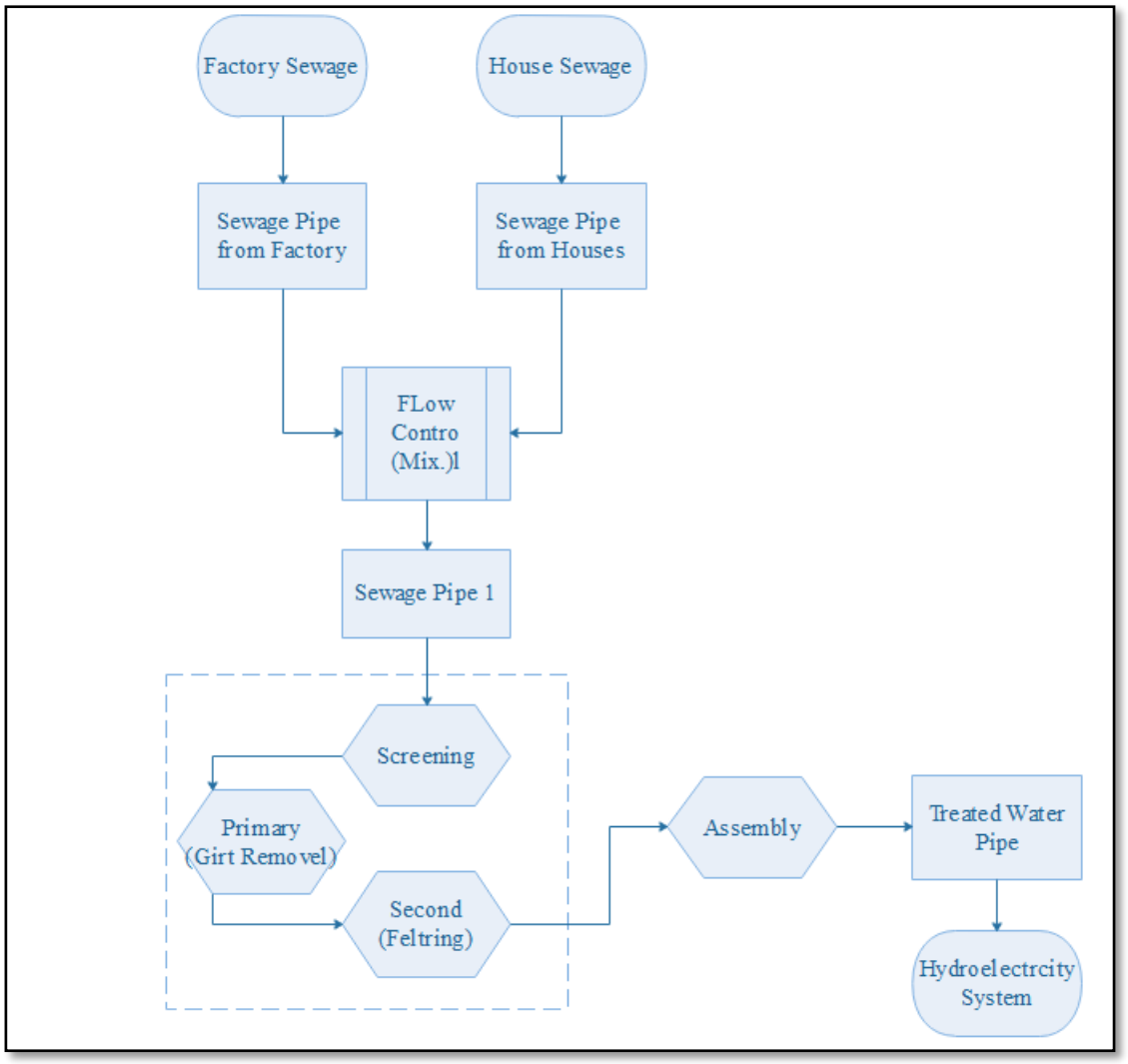


Figure 47 the Process Flow for the Sewage System

5.10 Housing Compound

The final infrastructure in the system is the housing complex, displayed in Figure 48. The Houses rely on three main Critical Infrastructures, the Water System, the Electricity System and the Sewage System. Both the water and the electricity are provider to the Houses and the waste is disposed to the Sewage System. Moreover, the Houses contain 27 different components, which are presented in Table 12.

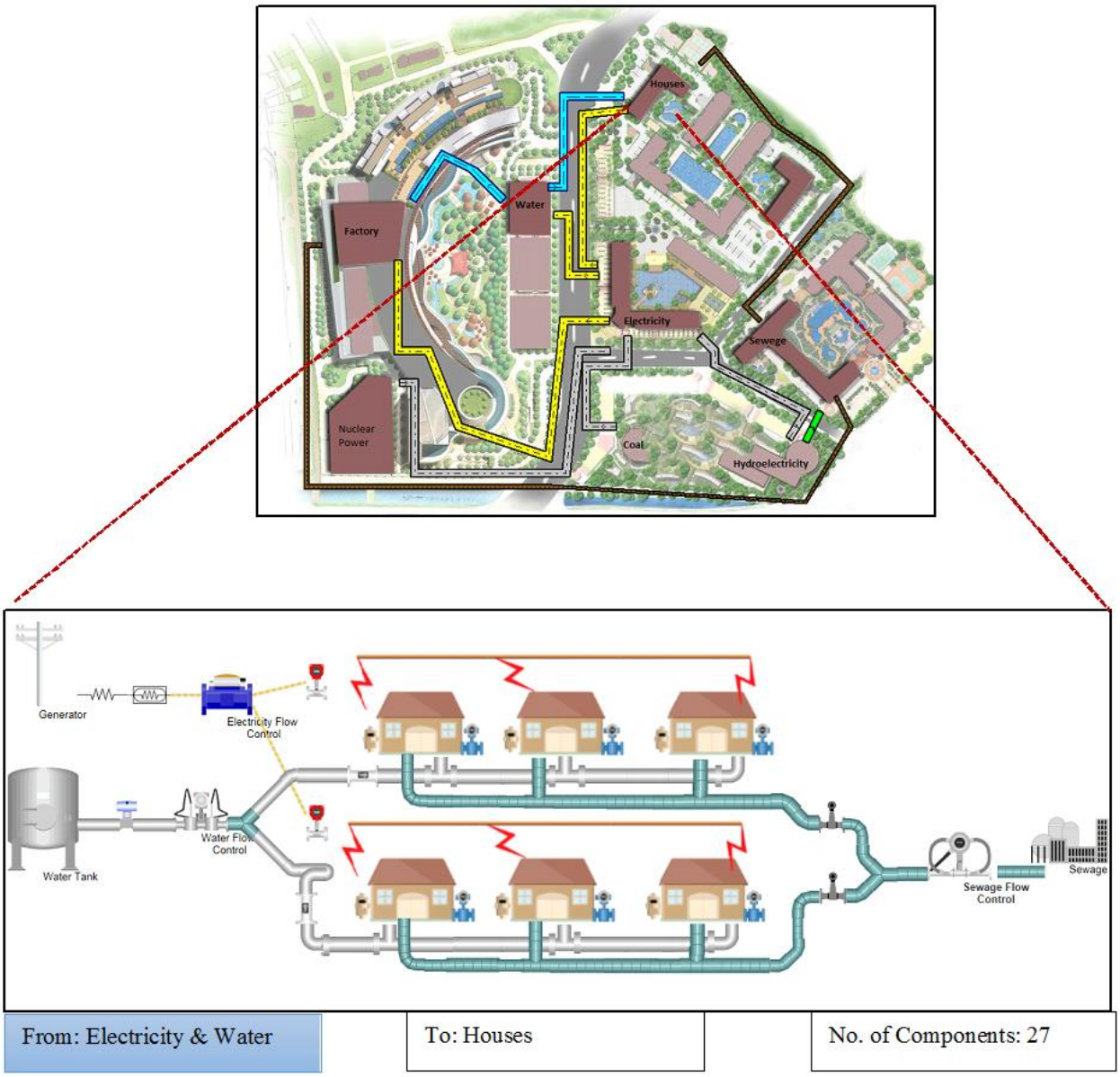


Figure 48 Houses System Components Diagram and location within the 8 Critical Infrastructures

Table 12: Simulation Houses Components

Components		
1. main cable electricity houses	2. cable electricity 6	3. pipe sewage 1
4. main pipe water houses	5. water meter 1	6. pipe sewage 3

7. electricity meter 1	8. water meter 2	9. pipe sewage 5
10. electricity meter 2	11. water pipe 1 in odd houses	12. pipe sewage 2
13. electricity cable1	14. pipe water 3 in odd houses	15. pipe sewage 4
16. cable electricity 3	17. pipe water 5 in odd houses	18. pipe sewage 6
19. cable electricity 5	20. water pipe 2 in even houses	21. Assembly sewage for odd houses
22. cable electricity 2	23. water pipe inside houses 1	24. assembly sewage for even houses
25. cable electricity 4	26. water pipe inside house 2	27. main sewage pipe

The housing compound, Figure 49, is an end-user of the services provided by the other infrastructure in the system. The diagram can be broken down into four main groups of objects. The groups include supply lines from other infrastructures, for example:

1. Electricity power plant input to houses

2. Water Distribution System inputs to houses

3. Housing compound

4. Sewage drains from houses to external infrastructure

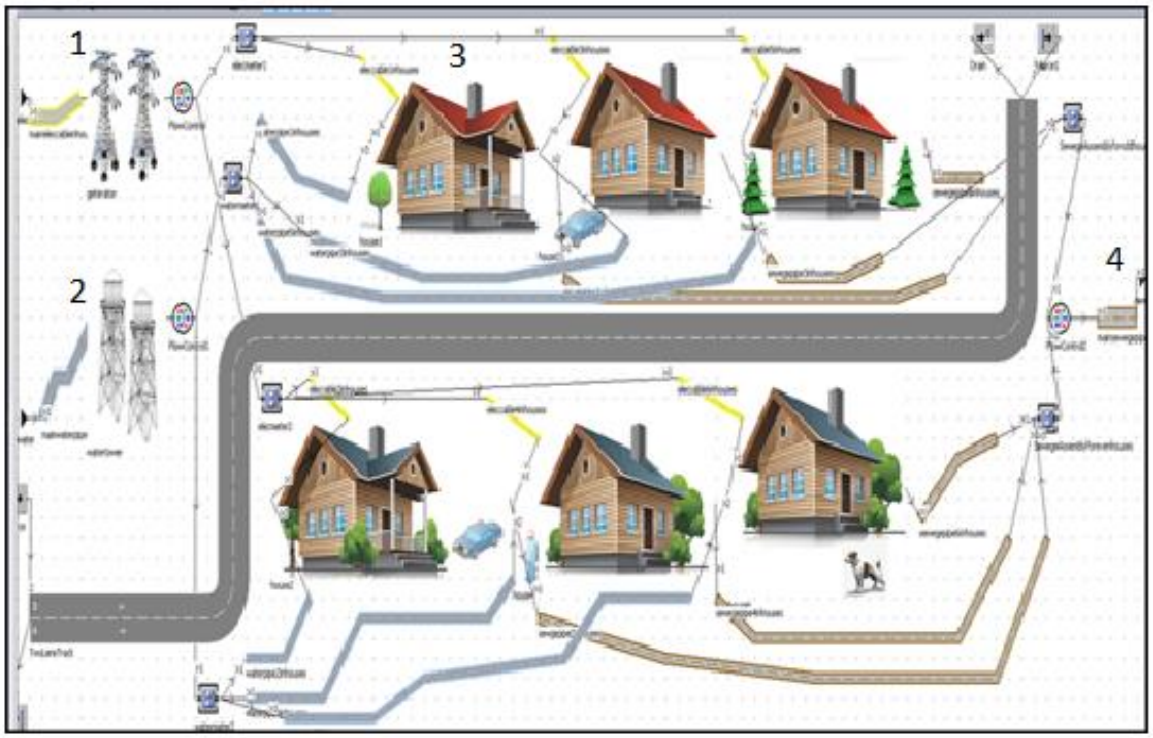


Figure 49 Simulation Plant for the Houses

Figure 50 indicates the functionality of the two parallel processes; the Water System and the Electricity System that are working within the Houses.

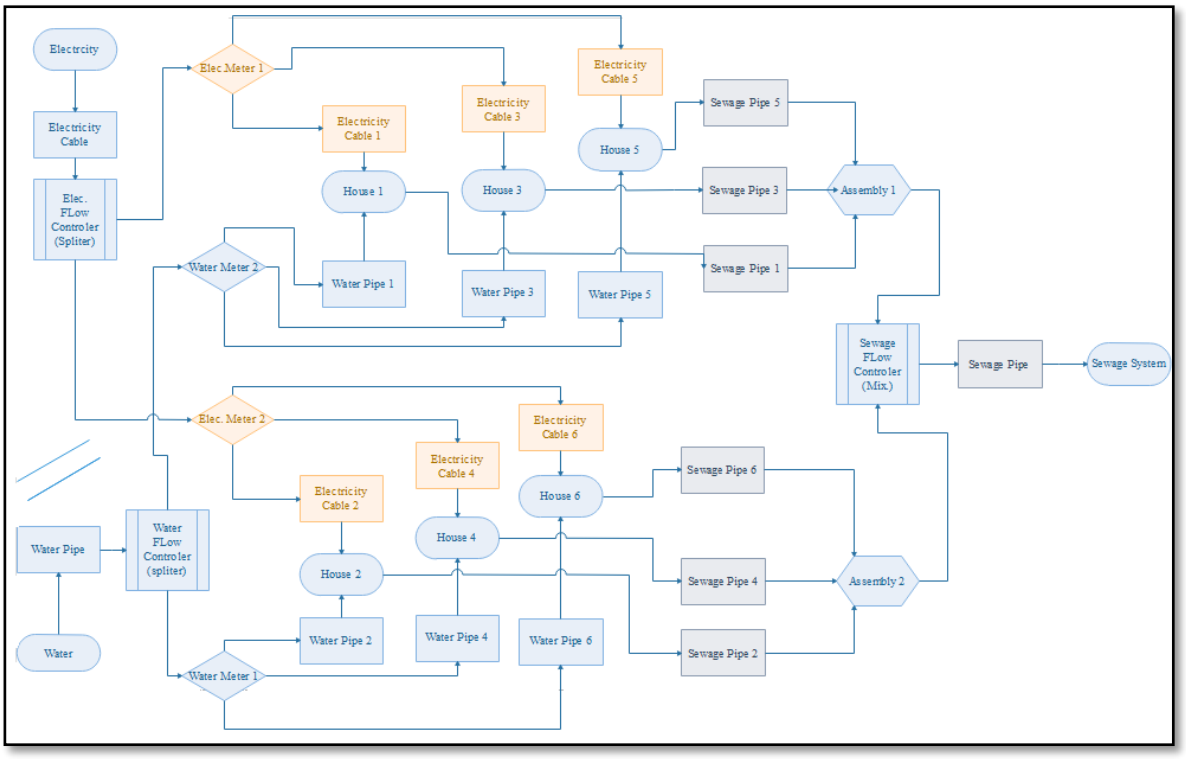


Figure 50 the Process Flow for the Houses

5.11 Simulation

To this point, we presented a digital model; in order to represent a realistic environment, two main steps need to be specified in the model to reach this aim. These include an event controller and the system availability percentage.

5.11.1 Event Controller

Firstly, an event controller was inserted into the simulation process to control speed was implemented. Figure 51 indicates the event controller interface used to influence the speed of operation. In addition, a set time for simulation can be inserted. For the dataset used in this research, we set the time for 2 days.

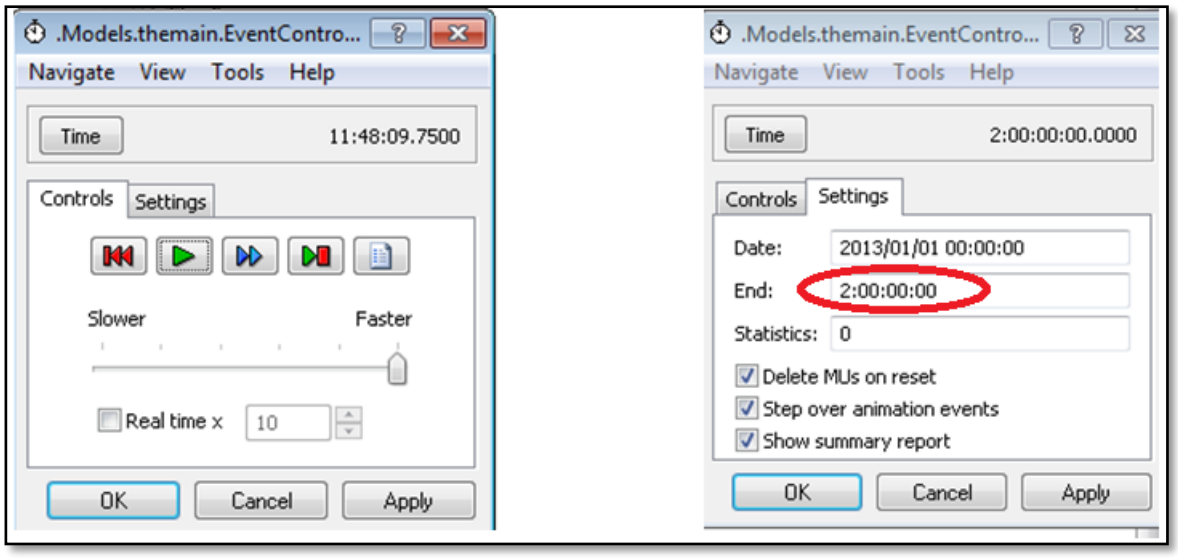


Figure 51 the Simulation Control Events

5.11.2 System Availability

Within the simulation, each of the components has a random failure implemented. This was done to make the system behave differently each time it runs and to account for faults, which occur in real-life Critical Infrastructures. Random failures are implement an Availability Percentage. The Availability Percentage refers to the chances of a machine or components being ready to use at any given time taking into account failures and blockages, which is calculated using the formula:

$$Availability = A/(A+B) \quad (1)$$

Where B represents MTTR, which is the Mean Time To Repair and A represents MTBF, which is the Mean Time Between Failures. In the following subsection, an example of abnormal behaviour, which could occur in the Water Distribution System and the Electricity Power Plant, is highlighted.

The behaviour changes are created by altering the availability percentages for the different components.

5.12 The Simulation Environment

Each object in our model has their own parameters that define how the materials flow through our environment. Figure 52 indicates the system functionality; the blue discrete block represents the water flow. While the green discrete blocks represent the electricity flow. The functionality of the system is stable and regular during run-time.

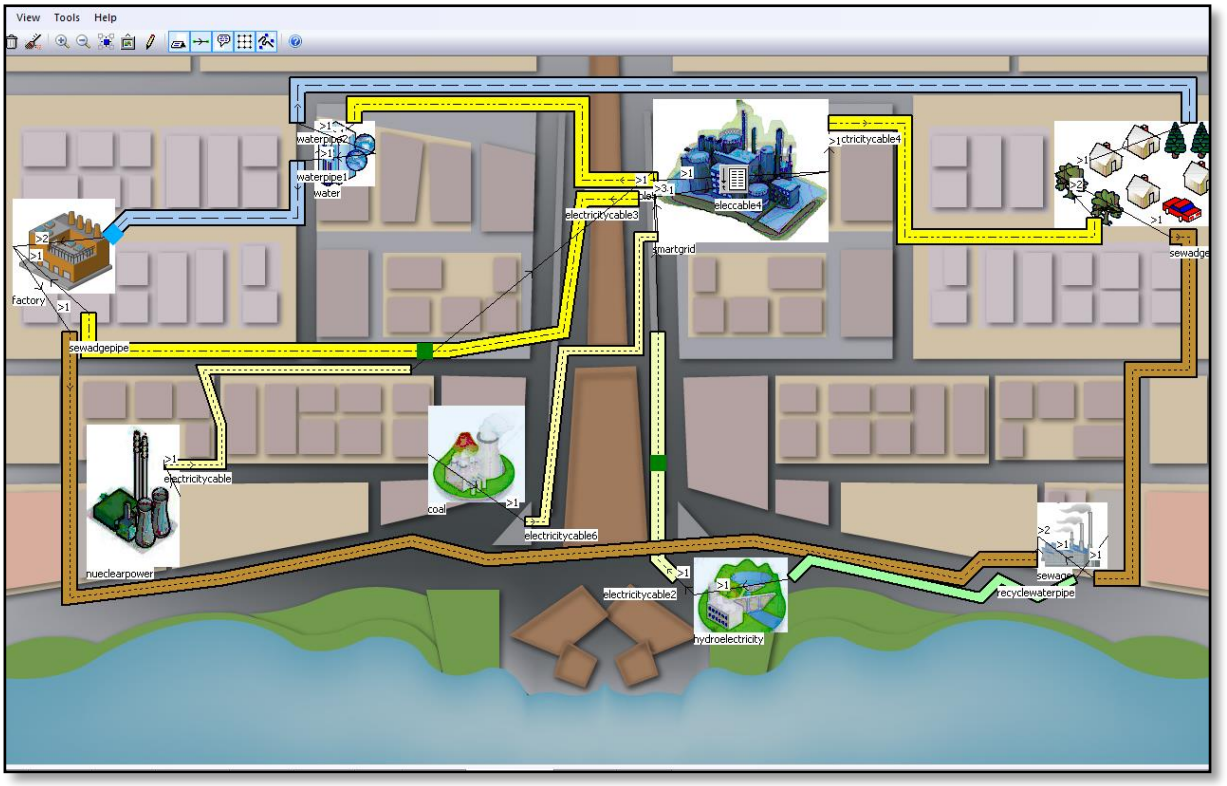


Figure 52 the System Flow Materials: Water, Electricity

In order to track the flow of services in the system, a Sankey Diagram is employed. The diagram visualises the material flow; moreover, it helps in producing a productive environment. Figures 53

to 55 indicate the electricity flow between the different materials and the methods. The heavier the line the more material is transported across the machines.

It is clear that the electricity flow, which is presented by the yellow line in Figure 53, is the heaviest between the Hydroelectricity System and the Electricity System, followed by three Critical Infrastructures, the Nuclear Power System, the Houses and the Water Distribution System. On the other hand, the lowest transportation is between the Electricity System, the Factory and the Coal System.

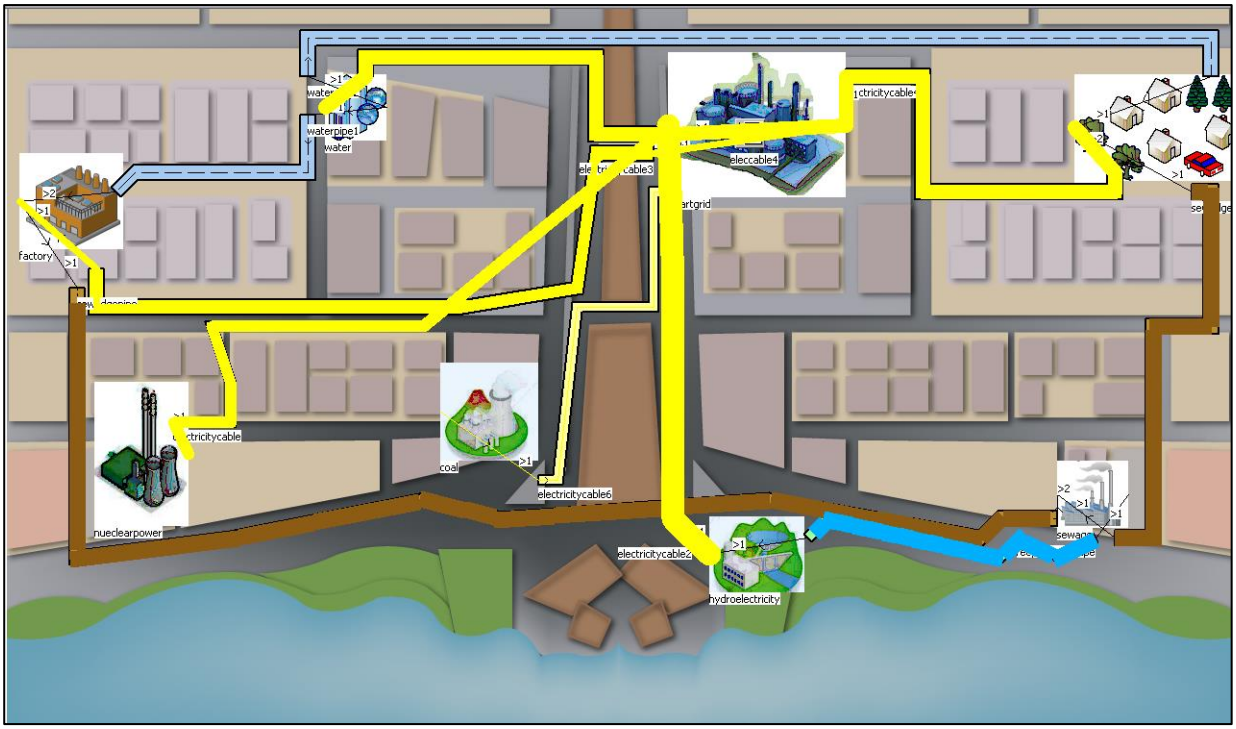


Figure 53 the Electricity transportation Simulation Sankey Stream between the Critical Infrastructures

Figure 54 provides more visualisation details regarding the level of material that flows in the Electricity System and it can indicate that the transportation from the Coal System was the Lowest. Finally, the electricity flow between the houses is almost equal and it was governed by the FlowControl component.

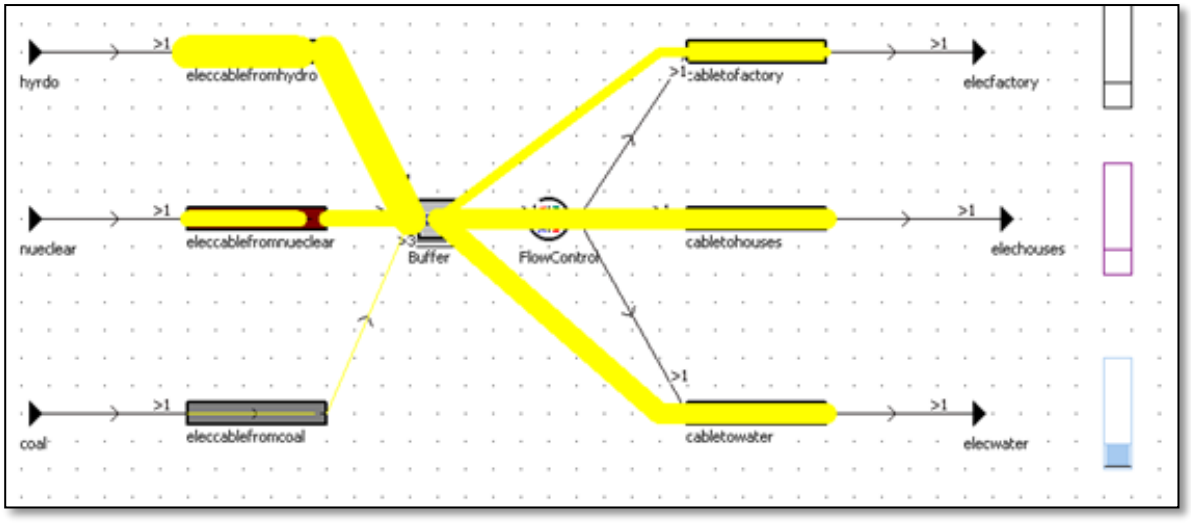


Figure 54 the Electricity transportation Simulation Sankey Stream within the Electricity System

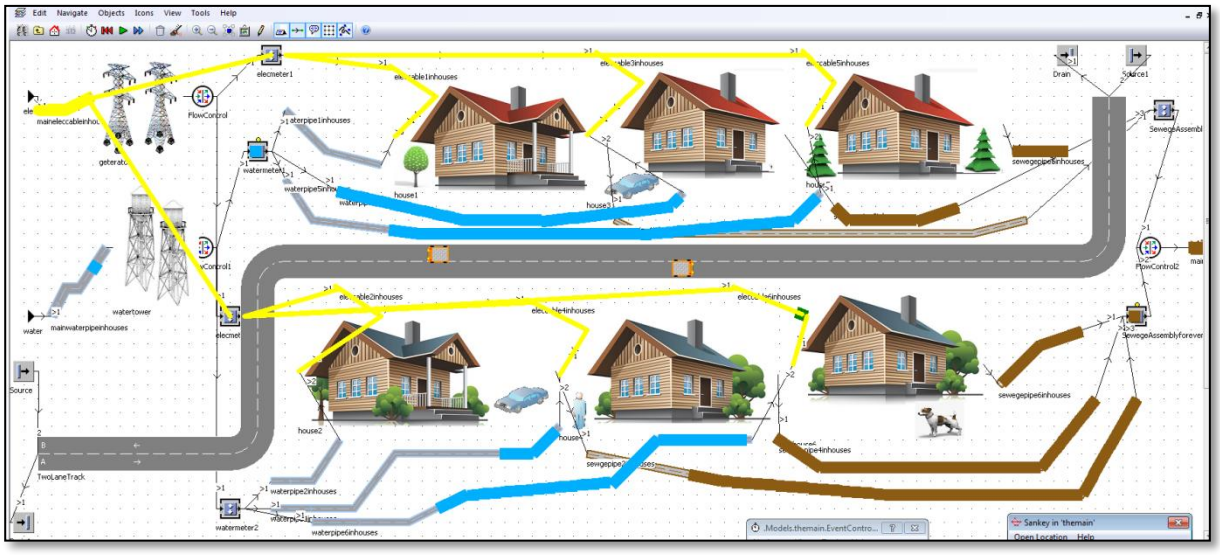


Figure 55 the Electricity transportation Simulation Sankey Stream between the Houses

Figures 56 to 58 indicates the water flow between the different materials and the methods. The heavier the line the more material is transported across the machines. It is clear that the water flow, which is presented by the blue line in Figure 56, was the heaviest between the Hydroelectricity System and the Sewage System; followed by Houses and the Factory Critical Infrastructures, which is presented also by Figure 57.

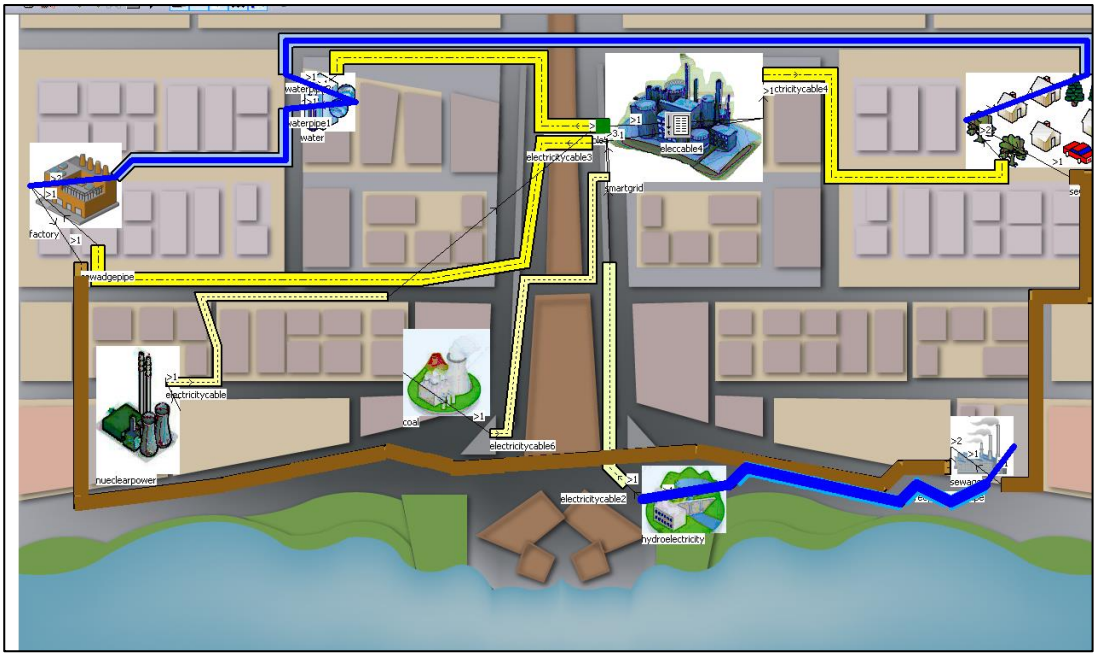


Figure 56 the Water transportation Simulation Sankey Stream between the Critical Infrastructures

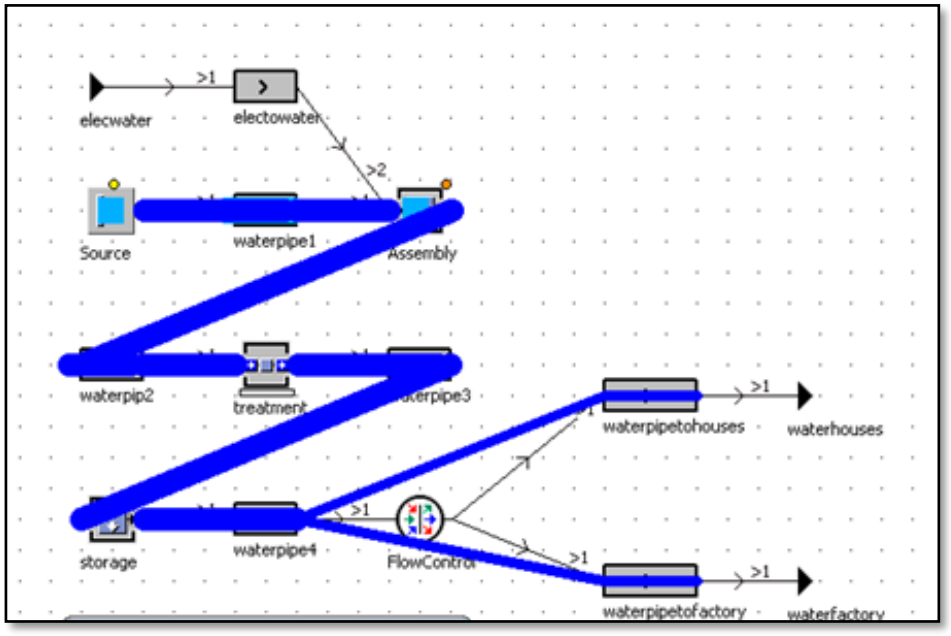


Figure 57 the Water transportation Simulation Sankey Stream in the Water System

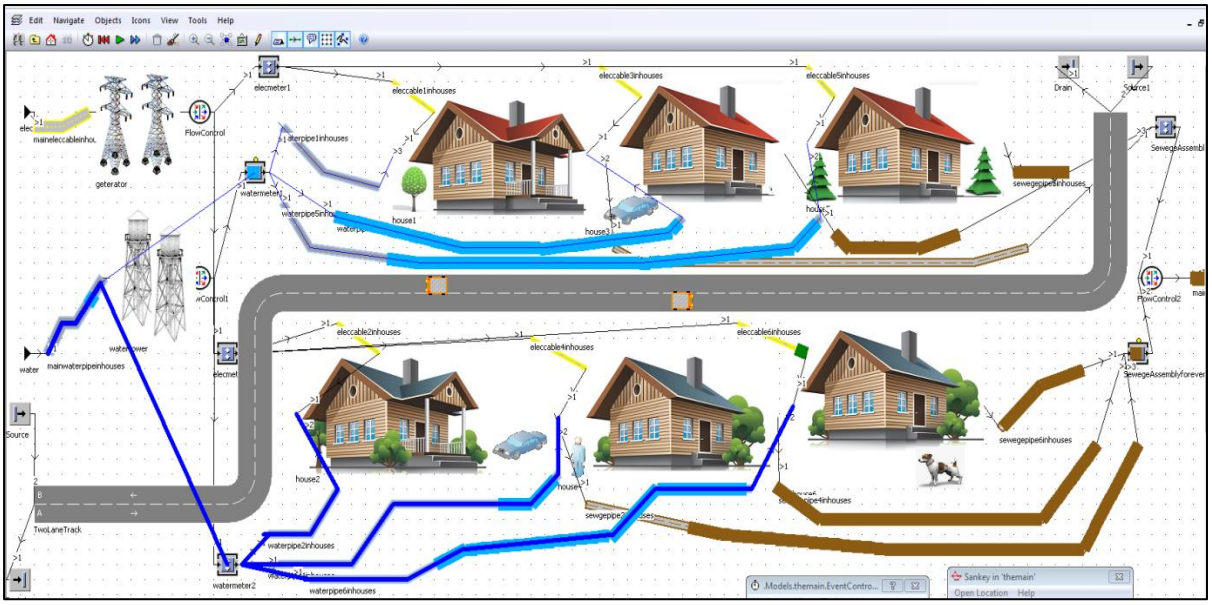


Figure 58 the Water transportation Simulation Sankey Stream between the Houses

Finally, Figure 58 presented the water transportation flow between the houses and it is clear that the levels of the materials are different.

5.13 System Behaviour

The data construction process depends on the connectivity between the different systems and the systems' faults, each based on realistic infrastructure behaviour. The expectation is that, through analysing the data from different attack scenarios on the system, cyber-attacks can be communicated between different infrastructures and suitable countermeasures can be established.

The trends in data patterns for both normal and abnormal behaviour can be identified and communicated to prevent future impacts. In this section, we present a snapshot for of normal and abnormal data. In order to understand the behaviour of the system, two datasets are constructed from The Water Distribution Infrastructure System. A normal system set constructed from a two days simulation. Then faults were introduced to the system as abnormal behaviours in order to construct a dataset of the system under attack. Blocks of data are extracted to prevent data overload

and to support building the features from both normal and abnormal datasets. This resulted in 691,200 rows of both normal and abnormal raw data for the two days of simulation. Based on the data collected, a number of features were extracted from both normal and abnormal behaviours.

The whole simulation consists of 147 components in total. However, Table 13 presents the Water Distribution components in detail. The numbers in the tables represent the units, which flow in the water pipe. However, more datasets are presented in the Appendix A that show how it helped shape the research.

Table 13: Component Description for Water Distribution Infrastructure

Abbreviation	Component Description
F1	Electricity cable from the Electricity Grid to the WD in the WD
F2	Water pipe 1 from the source in the WD
F3	WD Assembly
F4	Water pipe 2 in the WD
F5	WD treatment
F6	Water pipe 3 after the treatment in the WD
F7	WD Storage
F8	Water pipe 4 in the WD
F9	Water pipe from WD to Houses in the WD
F10	Water pipe from WD to Factory in the WD

5.13.1 Normal Data Collection

Table 14 shows data for normal behaviour for the Water Distribution System.

Table 14: Snap of the Model of the Normal Data Sample in the Water Distribution System

Time	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Normal Data Set										
01:57.0	0	3	1	0	0	0	0	1	0	0
01:57.2	0	3	1	0	0	0	0	0	1	0
01:57.5	0	3	1	0	0	0	0	0	1	0
01:57.8	0	3	1	0	0	0	0	0	1	0
01:58.0	0	3	1	0	0	0	0	0	1	0
01:58.3	0	3	1	0	0	0	0	0	1	0
01:58.5	0	3	1	0	0	0	0	0	1	0
01:58.7	0	3	1	0	0	0	0	0	1	0
01:59.0	0	3	1	0	0	0	0	0	1	0

5.13.2 Abnormal Data Collection

A number of recognised faults were introduced into the simulation in order to understand the different records that the system would produce and how to differentiate these records from normal operation of the system. As for the normal system, these faults were simulated over a period of two days to create a balanced dataset. Table 15 presents the faults where the percentages indicate the level of unavailability.

Table 15: Number of Different Faults in the Critical Infrastructures

Faults					
The main system	Water pipe from the water D.S to the House compound 40%	Electricity cable from the Electricity Grid to the House compound 30%	Electricity cable from the Hydroelectricity system to the Electricity Grid 20%	Electricity cable from the nuclear Power system to the Electricity Grid 60%	Sewage Pipe from the House compound to the Sewage system 25%
Coal System	Heat tube in the coal boiler 40%	Steam pipe in the coal condenser 50%	Electricity cable in the coal to the smart grid 25%		
	Electricity cable in the coal to the smart grid 40%	Coal turbine 35%	Coal tunnel to the boiler 35%		
Electricity System	Electricity cable from hydroelectricity system 60%	Electricity cable to houses 50%			
	Electricity cable from nuclear power 70%	Electricity cable to water distribution 60%	Electricity cable to factory 80%		
	Electricity cable from coal 40%	Electricity cable from hydroelectricity system 50%	Electricity cable to houses 60%		
Factory	Single processing unit (proc) 50%	Water pipe in the factory 40%			
Water System	Water pipe 1 50%	Water pipe to houses 60%			

	Electricity cable 40%	Water pipe 1 60%	Water pipe to factory 70%		
	Electricity cable 60%	Water pipe to houses 50%			
Houses	Elec Meter 1 20%	Elec Meter 2 50%	Water pipe to the meters 40%	Main sewage pipe from the houses to the sewage system 30%	Elec Flow control in the electricit y 70/30
	Electricity cable to the meters 50%	Water Meter 1 30%	Water Meter 2 40%	Water Flow control in the water 80/20	Flow control before the sewage pipe 80/20
Hydroelectrici ty System	Water pipe from sewage system 50%	Electricity cable 30%			
	Water pipe 1 40%	Turbine 50%			
Sewage System	Sewage pipe from houses 60%	Sewage pipe1 40%			
	Sewage pipe from factory 50%	Screening 30%	Water sewage to the hydroelectricit y system 60%		
Nuclear System	Cool water pipe in the nuclear system 40%	Reactor in the control rods 35%	Nuclear turbine 50%	Electricit y cable from the nuclear to the grid 40%	
	Steam line in the nuclear power 30%	Steam line 2 from the turbine to the condenser nuclear 45%	Generator 60%	Electricit y cable from the nuclear to the grid 30%	Cool water pipe to the pump in the

					condense r 40%

However, for the purpose of this section, we choose to present one of the abnormal behaviours in the Water Distribution System. It is clear that between the time 1:57 to 1:59 the level of the water was increased, which is indicated by Table 16.

Table 16: Snap of the Model of the Abnormal Data Sample in the Water Distribution System

Time	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**F9**	**F10**
Abnormal Data Set										
01:57.0	0	3	1	2	1	0	0	0	0	0
01:57.2	0	2	1	2	1	1	0	0	0	0
01:57.5	0	3	1	2	1	1	0	0	0	0
01:57.8	0	3	1	2	1	1	0	0	0	0
01:58.0	0	3	1	2	1	1	0	0	0	0
01:58.3	0	3	1	1	1	2	0	0	0	0
01:58.5	0	3	1	1	1	2	0	0	0	0
01:58.7	0	3	1	1	1	2	0	0	0	0
01:59.0	0	3	1	1	1	2	0	0	0	0

5.14 Summary

This chapter presented our Critical Infrastructures simulation model used for data collection. The simulation is based on a semi-structured interview that helped to shape a real-world system using the Siemens Tecnomatix Plant Simulation program. Data flows are also highlighted for the water and the electricity materials in order to give a clear visualized point of view by using the Sankey

Diagram. Finally, two samples of datasets were pointed out to show the system behaviour with normal and abnormal data sets. This data is used in the following section for our evaluation and results.

CHAPTER 6

EVALUATION

6.1 Introduction

The Water Distribution System is one of the key utilities in the infrastructure grouping and is heavily relied upon by the general population [140]. As such, the results presented focus on the Water Distribution System in this evaluation. This chapter is divided into two main sections: the outcomes of a semi-structured interview used to shape the second section. Moving to the second section, describes the attacks scenario followed by the results and data sharing process.

6.2 The Semi-Structured Interview Statistical Report

The outcomes of the interview are presented in a statistical report. The problem was displayed in detail with the practical solution, which the Saudi Water Ministry choose. The semi-structured interview discussed the following points:

- The role that water distribution plays in our lives and how important the water infrastructure is.

- The failures that face the water infrastructure

- The different attacks the system faces and the process to solve it

- Interconnectivity between the water infrastructure and other infrastructures.

6.2.1 Interview Outcomes

The outcome of each question is analysed individually and highlighted as follows:

Question 1:

This question was asked in order to understand how important the water system is and try to relate it to research.

Answer 1:

The infrastructure of water networks consists of several basic elements: water connections with a counter, sub-lines, main lines, storage, pumping stations and treatment stations. These elements are always exposed to some faults for many reasons. The damage increases within the small elements, such as the house connection pipes. Therefore, the authorized stakeholders need to provide suitable maintenance teams that commensurate with the needs of each element of the water system.

Question 2:

The second question asked about the problem that the water distribution had and if these problems involve security aspects.

Answer 2:

The water distribution system faces technical problems that can be highlighted as follow:

- Leakage resulting from natural disasters as floods.
- There is no clear plan that outlines the path of the water pipes.
- Lack of remote control systems and the water network work mostly by rotation, which made it easy to hack it.
- Old water network and the difficulties of renewing it.
- Lack of professional trainers, which made it hard to do preventative maintenance.
- Lack of spare parts
- The lack of security units guarding the water tanks against tampering
- Still, the Saudi water distribution system uses the old water meters.

On the other hand, the attacks, which the water distribution system faced, can be highlighted in two main points:

- Entering the customer account and changing the usage rate either by an insider or by an outsider.
- The possibility of disabling the subscriber system, and remote operating systems

Question 3:

Different regulation could be found in different organization. Therefore, this question was concerned with tactics and regulation toward technical failures.

Answer 3:

New rules have been changed in order to improve the system. Use a backup system to keep the information in a safe place. Update the control system and export the physical network to simulation programs. Reconstruct the old water network to cooperate with people's needs. Moreover, contract with new expertise that can maintain the water system, improve it, and create a new check-up system concerned with house water tanks in order to indicate any manipulation.

Question 4:

This question was asked in order to identify the interconnectivity between the water infrastructure and the other infrastructures.

Answer 4:

The water distribution system is connected to different infrastructures, such as a sewage distribution system, electricity grid, telecoms, *etc*. in addition, more interdependency can be found depending on expansion within systems.

Question 5:

The final question was about a realistic scenario, which the water distribution face and the way they have handle it. This question helped in implementing our system.

Answer 5:

The engineer highlighted three real scenarios. However, for the aim of this research one of these scenarios was only chosen to implement. The next section describes the scenario.

6.2.2 The Scenario

Heavy rainfall led to displacement of the main water line, which connects the water source to the main water station in Makkah. In addition, three sub-lines were shifted from their location leaving a major leakage. Moreover, the water affected the electricity cables near them and the availability was only 60% for two days.

As soon as the admin in the Saudi ministry received the notification, a maintenance team were sent to lock the main water taps. A number of procedures were taken to address the situation and minimize the problem. Table 17 indicates the procedures and the time, which the maintenance team took in order to fix the problem.

Table 17 the Procedure of fixing one of the failures in the water system by the maintenance team and the time.

Procedure	Time
Close the main water seal valve	60 minutes
Determine the size of the damage	30 minutes
Determine the appropriate way to solve the problem	60 minutes
Indicates the required pieces of materials, equipment and workers	30 minutes

Prepare the chosen requirements and bring it to the work site	120 minutes
Repair the water lines	14.50 hours
Open the main water seal valve and run the system gradually	60 minutes

Depending on the procedure, which the maintenance team took and the time, the ratio of non-operation of the main pipe was calculated. This ratio was used to simulate the scenario with Tecnomatix. The percentage of water pipe defects is calculated using the formula:

$$\frac{Total\ Time}{Number\ of\ Prepared\ Days} \times 100 \quad (2)$$

Depending on the formula, the ratio of non-operation in the main water pipe is 60%. By stimulating this scenario this helped reach the aims of this thesis.

6.3 Water Distribution System: Case Study

Based on the semi-structured interview this section was formed. The Water Critical Infrastructure consists of the following:

- Water distribution networks, such as mainlines, sub lines and household connections.
- Pumping stations.

- Tanks, such as operational and a reservoir.

- Dams for storage.

- Treatment stations.

These components are all present in the simulation and can be a target for an attacker. The choice to study the Water Distribution System is based on some aspects, which are considered as follows:

1) The importance of water and the direct impact on human life.

2) Depending on the economic collapse, some cities cannot invest in developing or changing their old water system.

3) Protecting the Water Critical Infrastructure includes securing all the transportation parts, such as sources, treatment plants, pipes and the distribution system (Bearing in mind the various water sources that might be used: ground sources, rain, surface water, sea, *etc*.).

4) The Water Critical Infrastructure is considered to be one of the core Infrastructures and is comprised of three Infrastructures in itself: the drinking-water supply system, the sewage system and the stormwater runoff control systems [141].

5) There is a connection between energy and the water system, which created a cascade failure between the systems. Therefore, by moderating the water and wastewater usage there is a direct impact on energy provision.

Figure 59 shows the Water Distribution System components.

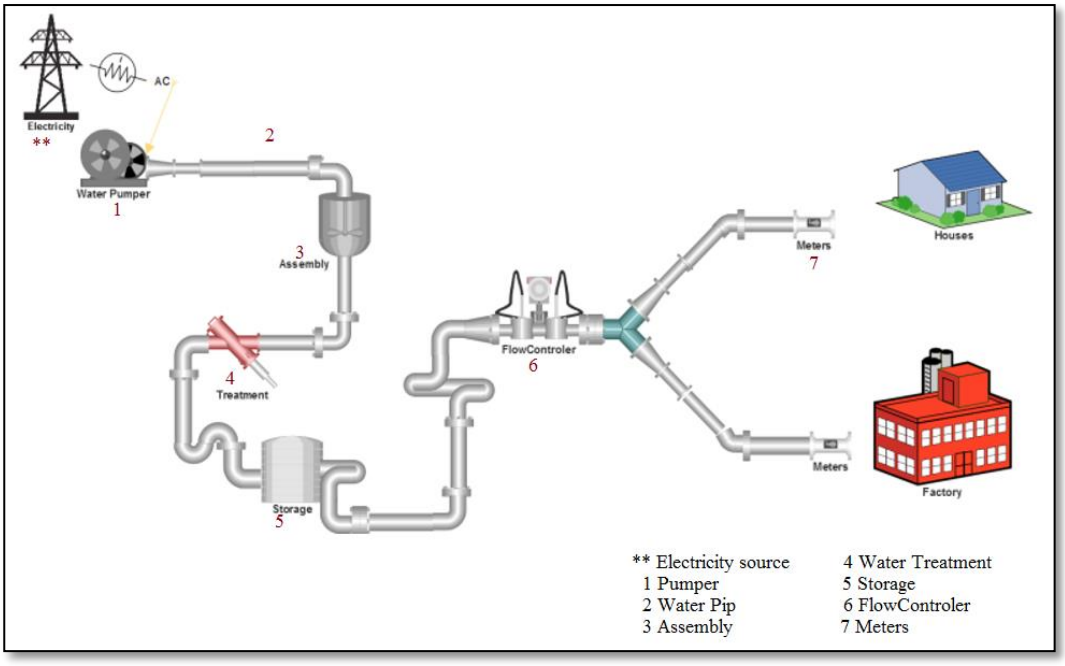


Figure 59 Water Distribution System Components Diagram

Starting with the water source, which needs to be available and accessible, water pipes and pumps are designed to meet the customer demands. One of the important components of this system is water treatment. This part needs to meet a standard depending on the location, in order to cover water consumption and quality. Water can be stored underground or in tanks and is used in peak demand. Finally, the distribution system that stores the water for houses or factory is used when needed. Figure 60 presents the Water Distribution System constructed in Tecnomatix Plant Simulation.

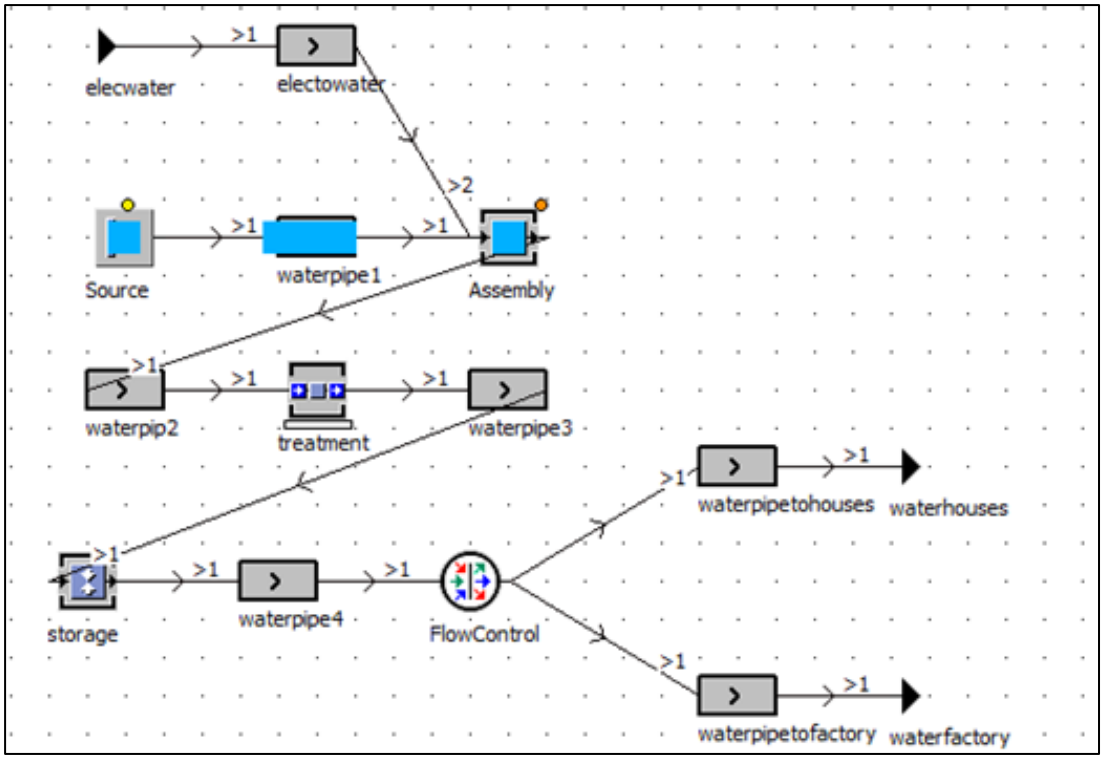


Figure 60 Plant Simulation for the Water Distribution System Components

6.3.1 The Abnormal Water Distribution Behaviours

A number of recognised faults are introduced to the Water System as abnormal behaviours, which are indicated in Table 18. One of these failures is chosen and explained in depth. As we highlighted earlier in the thesis, in order to present the productive simulation environment a '*system availability*' needs to be specified, which is clear in Table 18. Therefore, by reducing the availability percentage in certain components in the water system, a threat behaviour dataset can be constructed.

Table 18: The Water Distribution System Abnormal Behaviours

Failures	Components	Available	Not available
1	Electricity cable	60%	40%
	Water pipe to houses	50%	50%
2	Electricity cable	40%	60%
	Water pipe 1	60%	40%
	Water pipe to factory	70%	30%
3	Water pipe 1	50%	50%
	Water pipe to houses	60%	40%

As an example, water pipe 1 and the water pipe that links to the houses, only function correctly for 50% and 60% correspondingly during runtime. However, each time the availability percentage was reduced it was important to check the overall total system availability performance.

6.3.2 Data Figures

Figure 61 displays the abnormal flow in the Water System depending on the abnormal behaviour introduced to water pipe 1 and the water pipe connected to the houses, which is present in Table 17. A Bottleneck Analyser is used to show the statistical data graphically on top of each component. Figure 62 indicates a chart for the water abnormal behaviour effect on five components. It is clear that three components were blocked during processing. This included the assembly, treatment and the storage components.

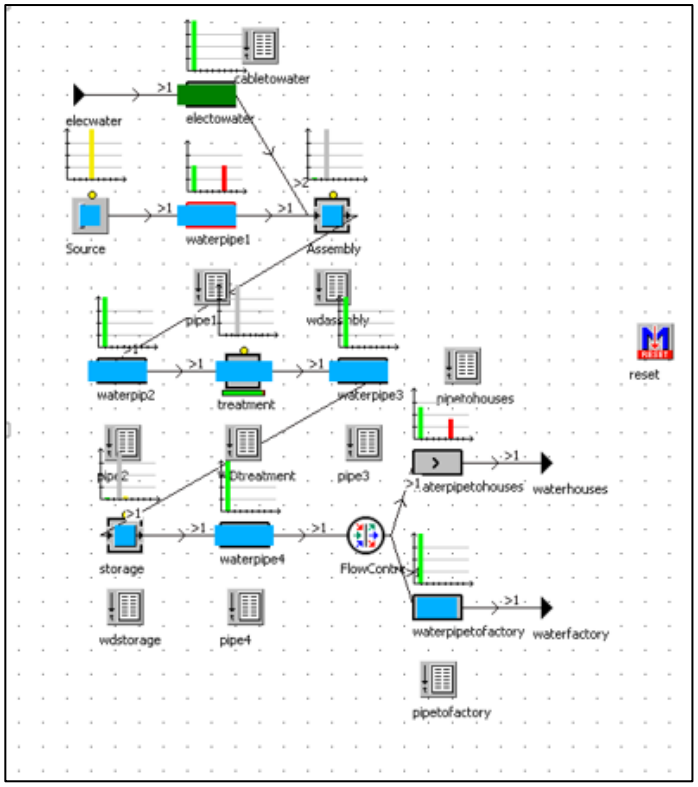


Figure 61 The Abnormal behaviour in Water Distribution System: water pipe 1 &water pipe to the houses

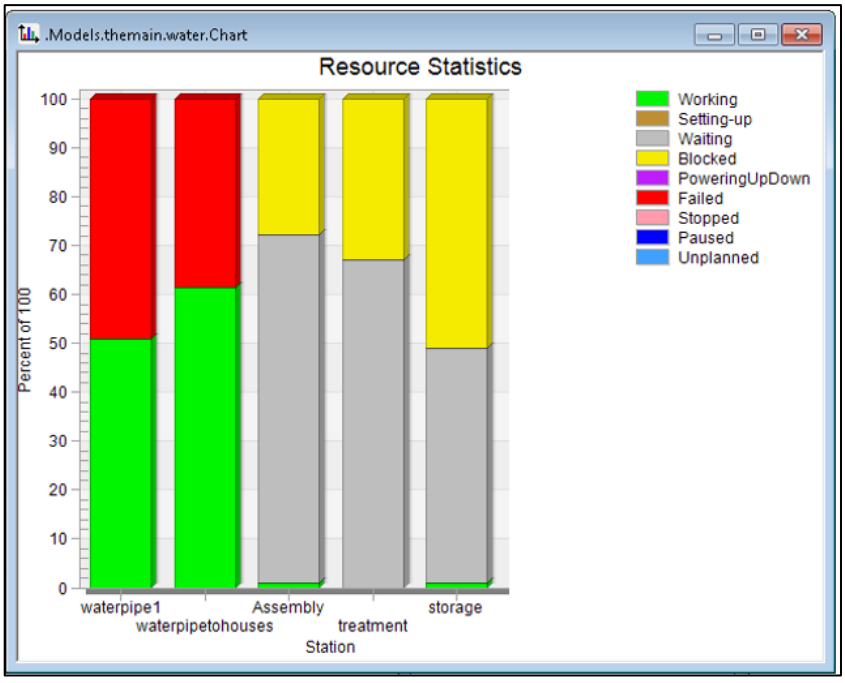


Figure 62 The Abnormal Chart for the Water Distribution System: water pipe 1 &water pipe to the houses

6.3.3 Datasets for Evaluation

This section presents the evaluation of the abnormal data set process for CIAIRS. CIAIRS uses a summative evaluation to improve the service within Critical Infrastructures. Data analysis is conducted using data visualisation to identify system anomalies and demonstrate that models of behaviour can be constructed and shared with other infrastructures.

Table 19 and 20 show data samples for normal and abnormal behaviour in the Water Distribution Infrastructure.

Table 19: The Water Distribution Normal Data Set

Time	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Normal Data Set										
01:51.5	0	3	1	0	1	0	0	0	0	0
01:51.8	0	3	1	0	1	0	0	0	0	0
01:52.0	0	3	1	0	0	1	0	0	0	0
01:52.3	0	3	1	0	0	1	0	0	0	0
01:52.5	0	3	1	0	0	1	0	0	0	0
01:52.7	0	3	1	0	0	1	0	0	0	0
01:53.0	0	3	1	0	0	1	0	0	0	0
01:53.2	0	3	1	0	0	1	0	0	0	0
01:53.5	0	3	1	0	0	1	0	0	0	0
01:53.8	0	3	1	0	0	1	0	0	0	0
01:54.0	0	3	1	0	0	0	1	0	0	0
01:54.3	0	3	1	0	0	0	1	0	0	0
01:54.5	0	3	1	0	0	0	1	0	0	0
01:54.8	0	3	1	0	0	0	1	0	0	0
01:55.0	0	3	1	0	0	0	0	1	0	0
01:55.3	0	3	1	0	0	0	0	1	0	0
01:55.5	0	3	1	0	0	0	0	1	0	0
01:55.8	0	3	1	0	0	0	0	1	0	0
01:56.0	0	3	1	0	0	0	0	1	0	0
01:56.3	0	3	1	0	0	0	0	1	0	0

01:56.5	0	3	1	0	0	0	0	1	0	0
01:56.7	0	3	1	0	0	0	0	1	0	0
01:57.0	0	3	1	0	0	0	0	1	0	0
01:57.2	0	3	1	0	0	0	0	0	1	0
01:57.5	0	3	1	0	0	0	0	0	1	0
01:57.8	0	3	1	0	0	0	0	0	1	0
01:58.0	0	3	1	0	0	0	0	0	1	0
01:58.3	0	3	1	0	0	0	0	0	1	0
01:58.5	0	3	1	0	0	0	0	0	1	0
01:58.7	0	3	1	0	0	0	0	0	1	0
01:59.0	0	3	1	0	0	0	0	0	1	0
01:59.3	0	3	1	0	0	0	0	0	0	0
01:59.5	0	3	1	0	0	0	0	0	0	0

Table 20: The Water Distribution Abnormal Data set

Time	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**F9**	**F10**
Abnormal Data Set										
01:51.5	2	3	0	0	0	0	0	0	0	0
01:51.8	2	3	0	0	0	0	0	0	0	0
01:52.0	2	3	0	0	0	0	0	0	0	0
01:52.3	2	3	0	0	0	0	0	0	0	0
01:52.5	2	3	0	0	0	0	0	0	0	0
01:52.7	2	3	0	0	0	0	0	0	0	0
01:53.0	2	3	0	0	0	0	0	0	0	0
01:53.2	2	2	1	0	0	0	0	0	0	0
01:53.5	2	3	1	0	0	0	0	0	0	0
01:53.8	2	3	1	0	0	0	0	0	0	0
01:54.0	2	3	1	0	0	0	0	0	0	0
01:54.3	2	2	1	1	0	0	0	0	0	0
01:54.5	2	3	1	1	0	0	0	0	0	0
01:54.8	2	3	1	1	0	0	0	0	0	0
01:55.0	2	3	1	1	0	0	0	0	0	0
01:55.3	1	2	1	2	0	0	0	0	0	0
01:55.5	1	3	1	2	0	0	0	0	0	0
01:55.8	1	3	1	2	0	0	0	0	0	0

01:56.0	1	3	1	2	0	0	0	0	0	0
01:56.3	0	2	1	2	1	0	0	0	0	0
01:56.5	0	3	1	2	1	0	0	0	0	0
01:56.7	0	3	1	2	1	0	0	0	0	0
01:57.0	0	3	1	2	1	0	0	0	0	0
01:57.2	0	2	1	2	1	1	0	0	0	0
01:57.5	0	3	1	2	1	1	0	0	0	0
01:57.8	0	3	1	2	1	1	0	0	0	0
01:58.0	0	3	1	2	1	1	0	0	0	0
01:58.3	0	3	1	1	1	2	0	0	0	0
01:58.5	0	3	1	1	1	2	0	0	0	0
01:58.7	0	3	1	1	1	2	0	0	0	0
01:59.0	0	3	1	1	1	2	0	0	0	0
01:59.3	0	3	1	0	1	2	1	0	0	0
01:59.5	0	3	1	0	1	2	1	0	0	0

The process involves detecting abnormal behaviour for sharing with other infrastructures. The Water Distribution Infrastructure consists of 10 components. Data collection was conducted with a sampling rate of 4 Hertz (which is one sample every 0.25 seconds) for two days. This sampling rate was chosen as often small, subtle changes of material flow take place within a 1 second period; a low sampling rate of 4 Hertz allowed us to detect these changes. This resulted in 691,200 rows of both normal and abnormal raw data after 2 days of simulation. From the datasets, records of data were constructed by selecting features from the datasets at 5-minute intervals [43]. The records used in the classification process consist of 572 normal and 572 abnormal behaviour records.

6.3.4 The Statistical Reports

Comparing Tables 21 and 22, it is clear that the production of the electricity in houses dropped from 71.75% to 61.19% and the production of the electricity in the factory also decreased by 10%.

The result indicates that the attacks, which accrued in the Water Distribution Infrastructure, have affected the production of the electricity in two other infrastructures including a factory and housing complex.

Table 21: The Normal Production Statistical Report for the Water Distribution Infrastructure

Object	Name	Mean Life Time	Production	Transport	Storage	Value added	Portion
nuclear power vapour	Steam vapour	25:48.4	7.81%	92.19%	0.00%	3.94%	
nuclear power vapour	water	41:52.2	3.85%	19.52%	76.63%	0.98%	
houses.house1.Drainelechouse1	electricity	01:51.8	71.75%	27.35%	0.89%	37.20%	
houses.house2.Drainelechouse2	electricity	01:51.8	71.75%	27.36%	0.89%	37.22%	
houses.house3.Drainelechouse3	electricity	4.2426	0.00%	100.00%	0.00%	0.00%	
houses.house4.Drainelechouse4	electricity	2.8284	0.00%	100.00%	0.00%	0.00%	
houses.house5.Drainelechouse5	electricity	2.8284	0.00%	100.00%	0.00%	0.00%	
houses.house6.Drainelechouse6	electricity	5.3852	0.00%	100.00%	0.00%	0.00%	
houses. Drain	car	01:42.2	0.00%	100.00%	0.00%	0.00%	
houses.Drain1	car	01:40.6	0.00%	100.00%	0.00%	0.00%	
factory. Drain	electricity	01:43.7	78.44%	20.60%	0.96%	41.16%	

Table 22: The Abnormal Production Statistical Report for the Water Distribution Infrastructure

Object	Name	Mean Life Time	Production	Transport	Storage	Value added	Portion
nuclear power vapour	Steam vapour	25:48.4	7.81%	92.19%	0.00%	3.94%	
nuclear power vapour	water	41:52.2	3.85%	19.52%	76.63%	0.98%	
houses.house1.Drainelechouse1	electricity	02:12.3	61.19%	38.05%	0.76%	31.73%	
houses.house2.Drainelechouse2	electricity	02:08.7	61.79%	37.43%	0.78%	32.06%	
houses.house3.Drainelechouse3	electricity	4.2426	0.00%	100.00%	0.00%	0.00%	
houses.house4.Drainelechouse4	electricity	2.8284	0.00%	100.00%	0.00%	0.00%	
houses.house5.Drainelechouse5	electricity	2.8284	0.00%	100.00%	0.00%	0.00%	
houses.house6.Drainelechouse6	electricity	2.8284	0.00%	100.00%	0.00%	0.00%	
houses. Drain	car	01:42.2	0.00%	100.00%	0.00%	0.00%	
houses.Drain1	car	01:40.6	0.00%	100.00%	0.00%	0.00%	
factory. Drain	electricity	01:59.9	68.27%	30.89%	0.83%	35.82%	

6.3.5 Descriptive Feature Extraction

In order to detect anomalous behaviour automatically for sharing with other infrastructures, features must be extracted from the dataset for use in the data analysis process. This subsection presents six features: Mean, Max, Min, Std, Mode and Median. Figure 63 illustrates the descriptive statistical types. The descriptive statistics are used in order to describe the data and not to conclude or prove a hypothesis [142]. The central aim is to help to describe the centre position for the data

and the common pattern. While, the measure of spread helps to study the distribution of the data, the two descriptive types are considered to be the most common data sets when analysing features that are used to understand the meaning of the whole data or a sample [143].

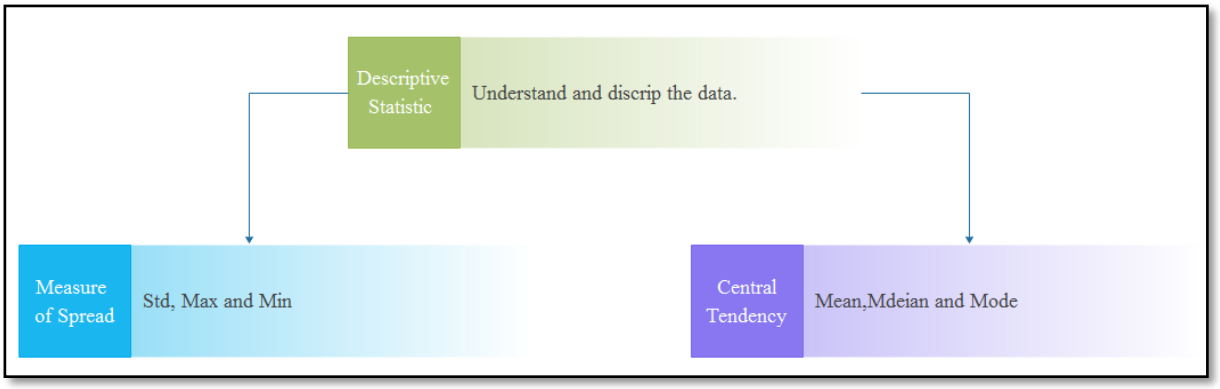


Figure 63 The Descriptive Statistical Types

Table 23 and 24 presents a sample of the feature vector records from the normal and abnormal behaviour in the water system constructed using the above features.

Table 23: Snap Shot for The Normal Water Distribution Vector Records

Normal	Avr. water pipe 1 from the source in the WD	max	min	mode	std	median	water pipe from WD to Houses in the WD	max	min	mode	std	median
Record 1	2.975	3	0	3	0.185482263	3	0.033333	1	0	0	0.179580334	0
Record 2	2.9941715	3	2	3	0.076153358	3	0.013322	1	0	0	0.114698316	0
Record 3	2.9975021	3	2	3	0.049937513	3	0.006661	1	0	0	0.081377265	0
Record 4	2.9941715	3	2	3	0.076153358	3	0.013322	1	0	0	0.114698316	0
Record 5	2.9975021	3	2	3	0.049937513	3	0.006661	1	0	0	0.081377265	0
Record 6	2.9941715	3	2	3	0.076153358	3	0.013322	1	0	0	0.114698316	0
Record 7	2.9975021	3	2	3	0.049937513	3	0.006661	1	0	0	0.081377265	0
Record 8	2.9941715	3	2	3	0.076153358	3	0.013322	1	0	0	0.114698316	0
Record 9	2.9975021	3	2	3	0.049937513	3	0.006661	1	0	0	0.081377265	0
Record 10	2.9941715	3	2	3	0.076153358	3	0.013322	1	0	0	0.114698316	0

Table 24: Snap Shot for the Abnormal Water Distribution Vector Records

Abnormal	Avr. water pipe 1 from the source in the WD	max	min	mode	std	median	water pipe from WD to Houses in the WD	max	min	mode	std	median
Record 1	2.9775	3	0	3	0.178943	3	0.033333	1	0	0	0.17958	0
Record 2	2.995004	3	2	3	0.070534	3	0.013322	1	0	0	0.114698	0
Record 3	2.995004	3	2	3	0.070534	3	0.006661	1	0	0	0.081377	0
Record 4	2.994172	3	2	3	0.076153	3	0.009992	1	0	0	0.099499	0
Record 5	2.999167	3	2	3	0.028855	3	0.009992	1	0	0	0.099499	0
Record 6	2.994172	3	2	3	0.076153	3	0.006661	1	0	0	0.081377	0
Record 7	2.997502	3	2	3	0.049938	3	0.006661	1	0	0	0.081377	0
Record 8	2.994172	3	2	3	0.076153	3	0.013322	1	0	0	0.114698	0
Record 9	2.769359	3	2	3	0.421419	3	0.006661	1	0	0	0.081377	0
Record 10	2.994172	3	2	3	0.076153	3	0.006661	1	0	0	0.081377	0

6.3.6 Detection Normal and Abnormal Behaviours

This subsection details the process involved in detecting abnormal behaviour for sharing with other infrastructures. The system consists of 147 components in total and for the purpose of the evaluation; nine components are used, which belong to the water distribution system.

The water system components provide 108 normal and abnormal features, and 48 records for the classification process consisting of 24 normal and 24 abnormal behaviour records. However, since the difference was clear in the mean and the standard deviation values, they were chosen to use as main features. Consequently, the nine water system components provide 18 features for each normal and abnormal behaviour records. Therefore, the water distribution vector records consist of 864 feature vectors, 432 normal and 432 abnormal behaviour records. The full normal and abnormal vector records are presented in Appendix C.

Table 25 presents the results of the classification process, which involved using seven well-known machine-learning algorithms. Fergus *et. al,*. clarifies their usage as follows [144], the linear discriminant classifier (LDC), the quadratic discriminant classifier (QDC), uncorrelated normal density based classifier (UDC), the polynomial classifier (POLYC), parzen classifier (PARZENC), k-Nearest Neighbour (KNNC) and the support vector classifier (SVC). LDC, QDC and UDC are

density-based classifiers. LDC generates discriminant functions for not normally distributed. While QDC assumes that the dataset follows a normal distribution function. UDC works parallel as QDC but computation of a quadratic classifier, between the classes in the dataset, is done by assuming normal densities with uncorrelated features [145].

The POLYC uses an untrained classifier with the dataset and adds polynomial features. PARZENC is considered a non-linear classifier that uses the training dataset and their parameters when building the classifier. The SVC support vector classifier is used in industry with a non-labelled dataset or few labels in order to point decide which class a new data point is in. NaivebC considers each of the contributed features independently to the probability [146].

Using the above classifiers to detect anomalous behaviour, ParzenC and KNNC obtain the best values. In addition, their Sensitivity and Specificity detection rates are also higher than other classifiers. This refers to the detection of normal and abnormal behaviours respectively.

Table 25: Classification Results for the Water Distribution System

Classifiers	AUC%	Sensitivity	Specificity
LDC	79.17	0.706	1.000
UDC	50.00	0.500	0.500
QDC	50.00	0.500	0.500
SVC	75.00	0.667	1.000
Parzenc	87.50	0.800	1.000
KNNC	87.50	0.800	1.000

Using the above techniques anomalous behaviour is detected and communicated to other infrastructures for mitigation and remediation planning, depending on the normal and abnormal

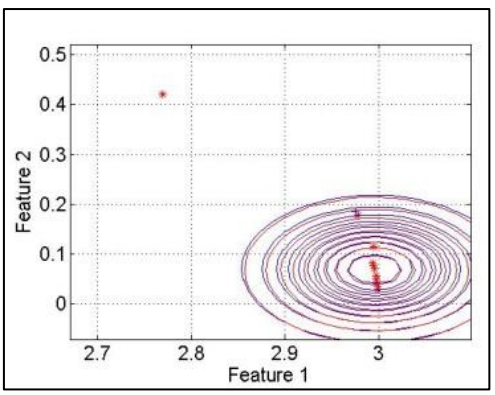
behaviour. Moreover, it is clear from the above techniques, that the selected classifiers can be used to identify abnormal behaviours in multiple Critical Infrastructure types.

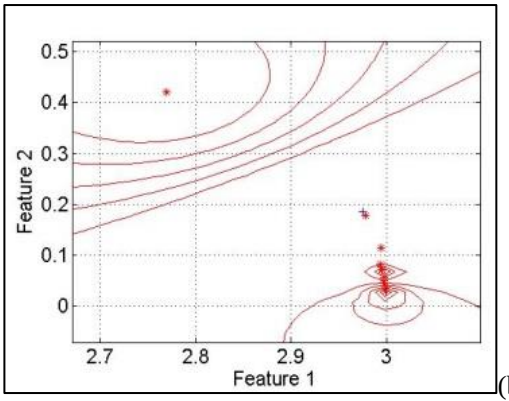
In addition, by using statistical measures or classification functions sensitivity and specificity offer positive and negative results for the data set, respectively. In other words, using these measures helped in identifying the level of accuracy detection between normal and abnormal system behaviour.

The sensitivity and specificity results were successful in detecting the normal behaviour. The results support that we could take advantage of assessing system behaviours to create a Critical Infrastructure support network. The aim is to alert other associated infrastructures about detected attacks, which could cause cascading impacts and advise an operator of the most appropriate response action.

Figure 64(a) displays a graph of the ParzenC classification. The ellipses displayed, refer to likelihood contours, where the points inside the ellipse are most likely to belong to that grouping. The blue ellipses consist of data that comes from the normal behaviour dataset and the red referring to threat behaviour data. Threat behaviour can be identified as a result of one grouping clearly standing out from the other.

The process functions by creating a scatter plot of the values from both of the selected features then drawing the ellipses based on the division of the data. The ellipses, displayed, refer to likelihood contours, where the points inside the ellipse are most likely to belong to that grouping.



(a)


(b)

Figure 64 ParzenC Visualisation & KNNC Visualisation

The blue ellipses consist of data that comes from the normal behaviour dataset and the red ones referring to threat behaviour data. Threat behaviour can be identified as a result of one grouping clearly standing out from the other. Similarly, Figure 64(b) displays a visualisation of the classification results for the KNNC classification process. Feature 1, on the x-axis, refers to one of the dominant features and Feature 2, on the y-axis, and refers to one of the less dominant features from the dataset. The graph displays that some changes in behaviour can be identified but often some are subtle and difficult to identify.

6.4 Summary

Critical Infrastructure interconnectivity is one the main challenges when countering the growing cyber-threat. This chapter presents a statistical report about the semi-structured interview by a professional Saudi engineer that works in the Water Saudi Ministry. Moreover, the chapter demonstrates how the system detects abnormal behaviour within a Critical Infrastructure and offers an approach for sharing the information with other infrastructures, using the human immune system as a reference model. A simulation approach was proposed for constructing big datasets for analysis. Using ParzenC and KNNC, two data classification techniques, we achieved high

accuracy in the detection of abnormal behaviours for the case study. Based on the results achieved in this chapter, we shaped the future work direction of the research.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 Introduction

This thesis presented a novel system, known as CIAIRS, for improving the level of security within Critical Infrastructures. The framework depends on a combination of unique techniques that helped in reaching the aim of the thesis. This chapter is considered to be the final stage of this research and the remainder of the chapter is structured as follows. Firstly, the thesis is summarised. Next, the main contributions of this work are presented. Finally, the future work, based on the research results, is discussed.

7.2 Thesis Summary

Designing an effective information system that helps in predicting abnormal behaviour in real time and shares the knowledge between different Critical Infrastructures was the main aim of this research. This section presents an overview of the thesis by summering each chapter, as follows:

- Chapter 1: In this chapter, we introduced the Critical Infrastructure research area. Presenting a brief about Critical Infrastructures and the challenges they face. Specifically, this included interdependency. Next, the motivation behind the research was highlighted.

After that, the aims and the objectives were indicated and finally, the novelty of the work was outlined.

- Chapter 2: This chapter presented an in-depth background on the Critical Infrastructure. In addition, a number of interdependency modelling examples were presented and the effects of cyber-attacks on the Critical Infrastructure network were highlighted.

- Chapter 3: the chapter highlighted different points regarding the Critical Infrastructures, such as the big data and the use of the analytics for predictive, descriptive, diagnosis and prescriptive analysis. In addition, the chapter indicates a pattern of communication that is similarly used in the framework, which is the publish-subscribe. This chapter also presented the case study for the eight Critical infrastructures, which were used to form the normal and the abnormal data. Next, the eight simulated Critical Infrastructures were detailed by presenting their components and the process flow for each. Moreover, the chapter details the data construction process through two examples: the electricity and water infrastructures. To present the flow of material Sankey diagrams were used. Finally, the chapter concluded with a snapshot of the normal and abnormal datasets for the water infrastructure. Finally, a review of the machine learning classes is highlighted. These points helped in clarifying the main aspect of the research.

- Chapter 4: This chapter presented the design of our system and framework; Critical Infrastructure Automated Immuno-Response System (CIAIRS). In addition, the location of the system within the ICS layers was outlined. Finally, a full description of each component was presented individually.

- Chapter 5: In this chapter the implementation on the Water System was presented. The chapter started by evaluating the normal and the abnormal datasets, which was collected from the Critical Infrastructure plant simulation. Next, the results were presented and the effectiveness of the CIAIRS system was evaluated.

- Chapter 6: this chapter presented the evaluation for the CIAIRS system.

- Chapter 7: A conclusion and suggestion of future work directions are presented in this current chapter., The chapter summarises the thesis findings and highlight the contribution of the research. Moreover, the chapter highlights some of the research limitations and the lessons learnt during the research.

7.3 The Research Contributions

This thesis contributes to the area of Critical Infrastructure security. Especially, it introduces a new way of thinking towards supporting Critical Infrastructure interdependency, in order to improve the level of security. By using our approach, the following is offered:

- The adaptive behaviour the system provides has increased the Critical Infrastructure level of security, especially by knowing the behaviour of the system patterns.

- We have developed the interconnectivity check process, in order to benefit from the connection between Critical Infrastructures. Using a dictionary to register the Critical Infrastructure interdependency has reduced the time it takes for the rest of the Critical Infrastructure to realise the attacks earlier.

- A novel communication process is used, which improves the level of security and performance for Critical Infrastructures and reduces the risk of cascading failures.

- CIAIRS uses a number of classifiers, which combine to give a better performance for the system during attack situations.

7.4 Summary of Thesis Findings

A number of challenges faced by Critical Infrastructures have been highlighted in this thesis, such as protecting SCADA. The research focused on interdependency between Critical Infrastructures. Consequently, the research investigated this challenge further using a semi-structured interview with a professional Saudi Engineer who works in the Saudi Ministry of Water. As a result of this investigation, the second stage presented a solution by using the interdependency between the Critical Infrastructures as a key to alarm the different systems. Therefore, a simulation of eight Critical Infrastructures was created in Tecnomatix and the results of the simulation helped in identify the behaviour patterns of the system. The patterns built up during the training helped to identify specific attacks.

Using seven classifiers, CIAIRS provides the operator with the ability to determine if the system is under attack or not. In addition, to prevent system overload, a window is used to extract the data. CIAIRS does not control the data it only spots specific types of information about attacks and shares it with interconnected infrastructures.

7.5 Research Limitations

Although the research achieved its aim and objectives, there were some unavoidable limitations. First, Adequacy of the Sample data: Depending on the nature of the research problem, the sample may be considered small and by enlarging the sample size that could generate more results that are accurate. Furthermore, if the sample is magnified this gives us the ability to use more complex models and classifiers. However, given the nature of Critical Infrastructures and their security needs, obtaining large amounts of real data was not possible.

Second, Simulation: the complexity of a system design is difficult to assess using a simulation, primarily due to time factors and access to real-life data. While many orders of complexity can be described this will never be representative of actual functions and processes in real-world scenarious. For this reason a comprehensive assessment of the appraoch posited in this thesis was not possible. Third, Time Constraints: the research constructed several Critical Infrastructures using an industry recognised simulator – having additional time that would have allowed me to explore the tool in more detail and provide a much deeper design and implementation much more representative of real-life installations. This would have yield more data and uncovered more complex patterns of behaviour.

Finally, the design of the research interview was cross-sectional, which means the data were gathered at one specific point only without pre or post testing events. By including more specific questions in the semi-structured interview and testing it could have helped to address particular issues later in the study.

7.6 Future Work

By reaching this section the thesis has achieved the research aims, findings and novel contributions. However, these contributions have themselves raised some interesting questions. The results in this thesis propose a strong foundation for future work in three main sections, which can be highlighted as follows:

7.6.1 Broadcasting with Different Systems

Our proposed CIAIRS broadcasting system assumes that different information is broadcasted between different Critical Infrastructures to provide an early alarm mechanism. Moreover, the different infrastructures use the same CIAIRS plug-in system, which increases the level of confidentiality, integrity, facilitates communication and the exchange of information.

However, CIAIRS can communicate with different plug-in systems (displayed as *com* in Figure 65). It would be interesting to see how CIAIRS can support broadcasting to different types of existing security systems. By integrating this method, this will save costs, as existing technology can be used.

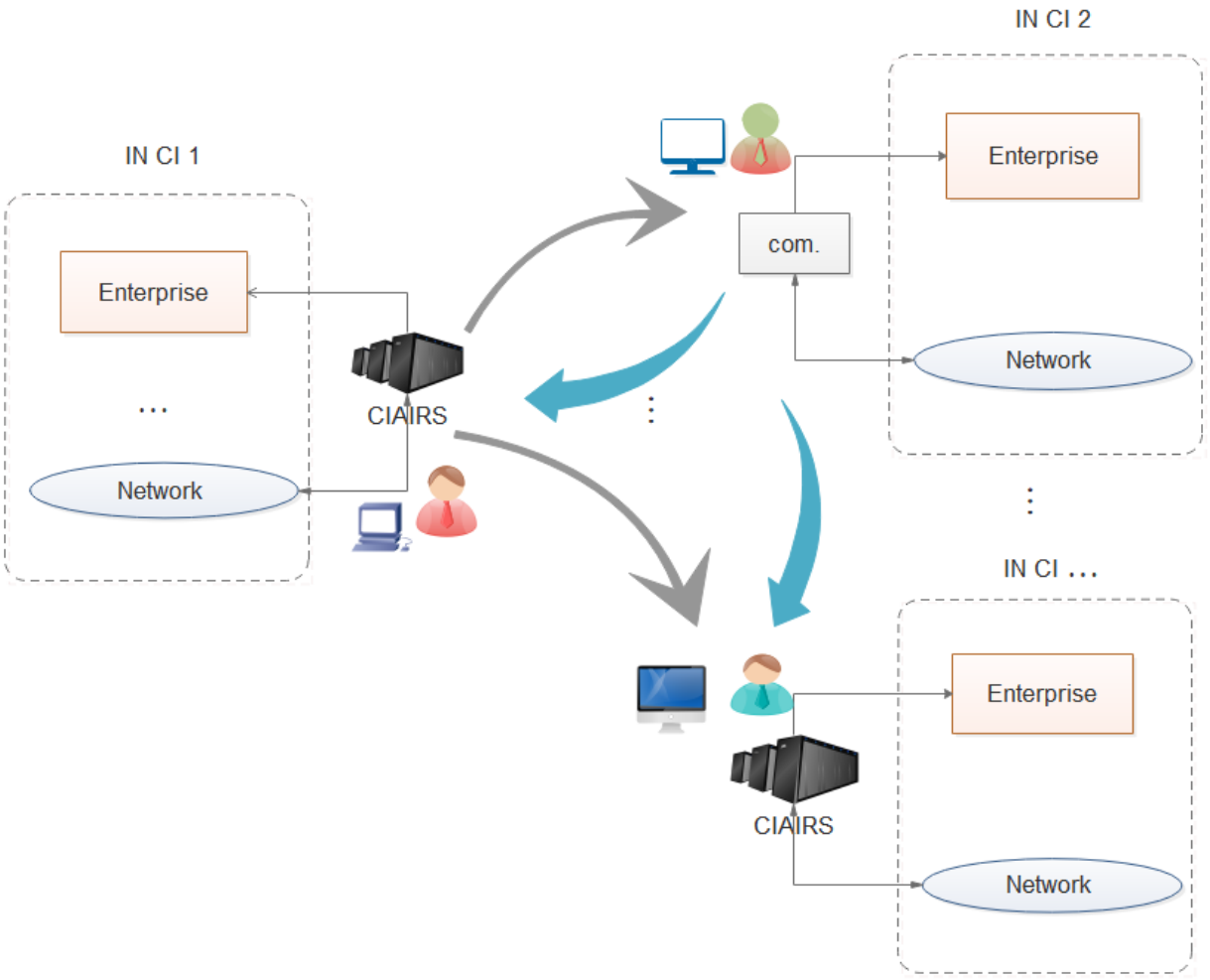


Figure 65 the Future CIAIRS Broadcasting Process

7.6.2 Automated Security Key

Another assumption that was proposed in this thesis, is that the CIAIRS broadcasting system uses an operator in order to ensure a certain level of protection, confidentiality and integrity before the broadcasting process. In addition, checking and auditing any information before sending it to the other Critical Infrastructures is governed by the human operator.

Figure 66 indicates how this process can function using an automated security key instead. For this, we envision that an authentication protocol can be used in a future design, in order to create a confidential channel between the different systems. This protocol allows the different Critical

Infrastructures to communicate and share the different information over a non-secure network by using “tickets” to prove their identity. Although, the use of a protocol ensures a high level of security during the broadcasting scheme, this strategy needs more investigation in order to evaluate the level of efficiency toward our system.

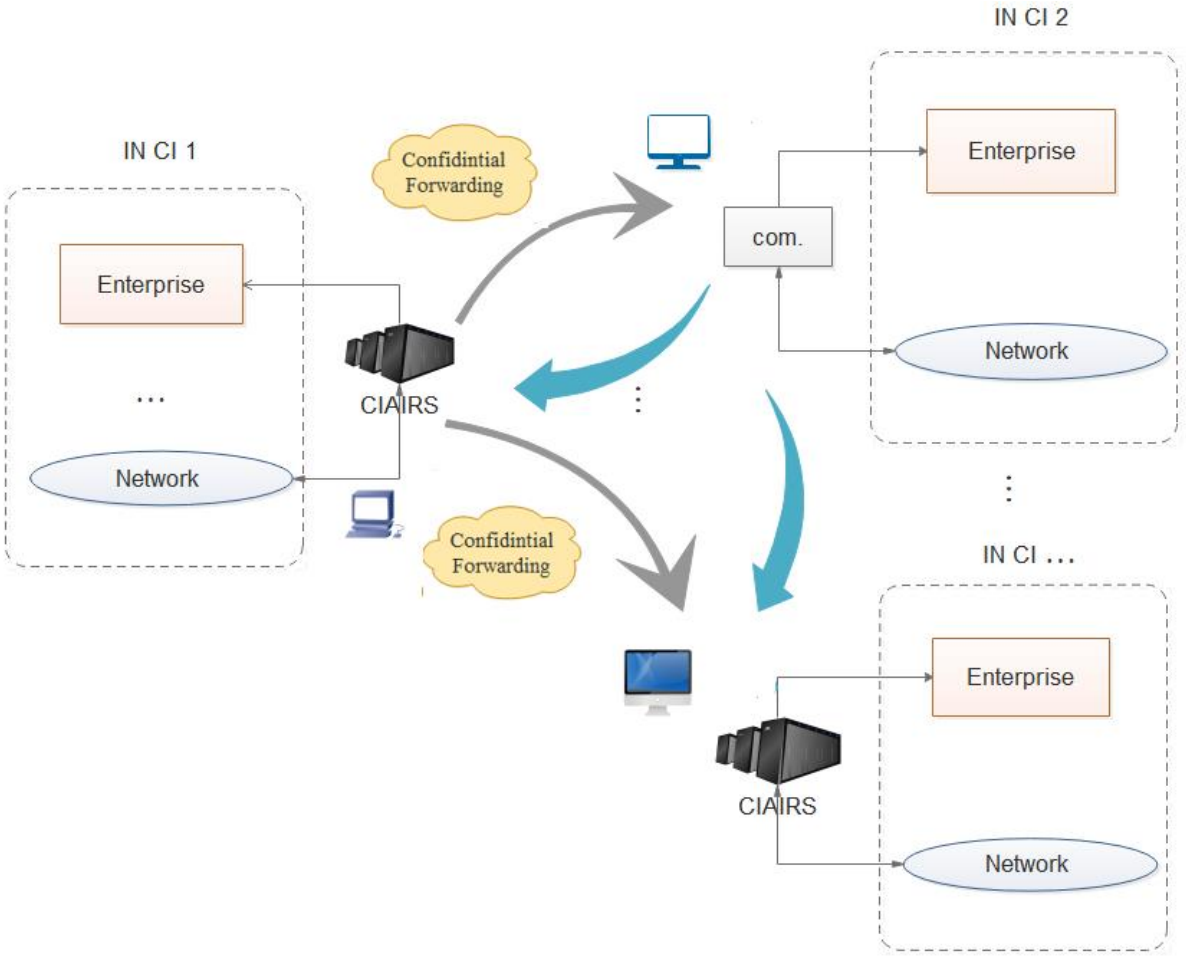


Figure 66 the Future CIAIRS Automated Security Channel

7.6.3 International Interdependency Support

Although the simulation was modelled on a real system, it is possible to expand the approach and construct additional infrastructures in order to assess the effect of international interconnectivity.

This would be an ideal expansion of the study. The benefit of using international interdependencies would mean that the effects of service provision across borders could be assessed.

It could be interesting to consider this model with different importance, for example, depending on their size or their international support. This model would help with understanding very complex problems, and potentially offer global recommendations. However, we need to bear in mind the cost, the different technical approaches and regulation measures in each country to be able to expand the simulation model.

7.6.4 CIAIRS in a Real Experiment

This research has measured the performance of the CIAIRS framework through simulation. The simulation helped in evaluating the performance of CIAIRS; but this approach will not replace a reality investigation. Therefore, it would be a great benefit to consider the effectiveness of the system using real-world critical infrastructure data. This would not only test the effectiveness of the system, but it would act as a test for the simulation environment to support its realism.

Just to emphasise how important this point could be a short phone interview was made with one of the Project Managers in the Ministry of Water and Electricity in Jeddah, Saudi Arabia. The interview focused on understanding the different attacks that the system faces and how they are handled. Moreover, what are the problems they are facing and if the interdependency in the water system is clear. Depending on the outcomes of the semi-structured interview, considering this point as a future work would give the system more credibility and would add more benefit especially for security researchers.

7.7 Concluding Remarks

Critical Infrastructures play a significant role in the world around us. Their service provision has become more widespread, to the point where it is ubiquitous in many societies. Infrastructure is the main source of development and economic construction process of any country. The Critical Infrastructure complexity introduced by interconnectivity is one of the significant challenges faced by distributed systems. By supporting interconnected networks, the system helps prevent the spreading the cyber-attacks within Critical Infrastructure interconnected networks. It is important to protect the assets and maintain their conditions under any circumstances.

Due to this increased connectivity, now, more than ever before, Critical Infrastructures face a number of possible digital threats. As a result, CIP has become a significant topic for research focus.

Our research presents a technique for the detection of abnormal behaviour within Critical Infrastructures and offers an approach for sharing the information with other infrastructures, based on the human immune system as a reference model. In addition, the development and evaluation of CIAIRS is presented. In order to reach the aim of this research a highly comprehensive simulation program is developed using Siemens Tecnomatix. The simulation put forward can be used to create substantial datasets. Systems, such as CIAIRS can assist to counter the growing cyber-threats and the risk of cascading failures. The research indicates the CIAIRS's framework. The various components and mechanisms were highlighted in order to presents the role of the CIAIRS, which shares information with other infrastructures, to create a distributed support network for enhanced cyber-security. Finally, a new solution was presented in this research, known as CIAIRS, for sharing recommendations within a network of interconnected Critical Infrastructures.

REFERENCES

[1] C. M. Lawler and S. A. Szygenda, “Components of continuous IT availability & disaster tolerant computing,” *2007 IEEE Conf. Technol. Homel. Secur. Enhancing Crit. Infrastruct. Dependability*, pp. 101–106, 2007.

[2] E. A. Fischer, “Cybersecurity issues and challenges: In Brief,” *Cybersp. Threat Landsc. Overview, Response Authorities, Capab.*, p. 12, 2016.

[3] L. Montanari and L. Querzoni, “Critical Infrastructure Protection : Threats , Attacks and Countermeasures,” no. March, pp. 1–164, 2014.

[4] J. B. Camargo Jr. and L. F. Vismari, “Challenges in Safety Assessment of Complex Critical Infrastructures,” in *2011 Fifth Latin-American Symposium on Dependable Computing Workshops*, 2011, pp. 37–38.

[5] B. S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, “Identifying, Understanding, and Analyzing: Critical Infrastructure Interdependecies,” *Control Syst. IEEE*, vol. 21, no. 6, pp. 11–25, 2001.

[6] S. M. Rinaldi, “Modeling and Simulating Critical Infrastructures and Their Interdependencies,” in *System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on*, 2004, vol. 00, no. C, pp. 1–8.

[7] R. Gorvett, D. Ph, and N. Liu, “Interpretive Structural Modeling of Interactive Risks,” *Enterp. Risk Manag. Symp. Soc. Actuar.*, pp. 1–10, 2006.

[8] S. Walsh, S. Cherry, and L. Roybal, “Critical infrastructure modeling: An approach to characterizing interdependencies of complex networks & control systems,” in *2009 2nd Conference on Human System Interactions*, 2009, pp. 637–641.

[9] W. Hurst, M. Merabti, and P. Fergus, “A Survey of Critical Infrastructure Security,” in *Critical Infrastructure Protection VIII, A Survey of Critical Infrastructure Security ,8th IFIP WG 11.10 International Conference*, 2014.

[10] B. Hämmerli and A. Renda, “Protecting Critical Infrastructure in the EU,” Brussels, 2010.

[11] T. Denis, “Managing Communications in Critical Infrastructures Protection,” in *2010 Second International Conference on Computer Engineering and Applications*, 2010, pp. 11–15.

[12] S. D. Wolthusen, “GIS-based command and control infrastructure for critical infrastructure protection,” *Proc. - First IEEE Int. Work. Crit. Infrastruct. Prot. IWCIP 2005*, vol. 2005, pp. 40–47, 2005.

[13] National Institute of Standards and Technology, “Framework for Improving Critical Infrastructure Cybersecurity,” *Natl. Inst. Stand. Technol.*, vol. 1, pp. 1–41, 2014.

[14] J. McIntyre and M. Warren, “Protection of New Zealand in the age of Information Warfare,” *Proc. Prot. Infrastruct. Third Aust. Inf. Warf. Secur. Conf.*, no. October, pp. 235–240, 2002.

[15] D. Command and F. Leavenworth, “Critical Infrastructure Threats and Terrorism,” in *DCSINT handbook No. 1.02*, 1st ed., no. 1, Distribution Unlimited, 2006.

[16] F. S. Yusufovna, F. A. Alisherovich, M. Choi, E. Cho, F. T. Abdurashidovich, and T. Kim, “Research on Critical Infrastructures and Critical Information Infrastructures,” in *2009 Symposium on Bio-inspired Learning and Intelligent Systems for Security*, 2009, pp. 97–101.

[17] L. Baud and P. Bellot, “Securing a Critical Infrastructure,” in *2010 IEEE RIVF International Conference on Computing & Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, 2010, pp. 1–4.

[18] T. M. Wilson, C. Stewart, V. Sword-Daniels, G. S. Leonard, D. M. Johnston, J. W. Cole, J. Wardman, G. Wilson, and S. T. Barnard, “Volcanic ash impacts on critical infrastructure,” *Phys. Chem. Earth*, vol. 45–46, pp. 5–23, 2012.

[19] L. Loroch, P. Madrzycki, and M. Swiech, “The integrated airbase protection system,” *Proc. 3rd Int. Conf. Recent Adv. Sp. Technol. RAST 2007*, pp. 119–122, 2007.

[20] P. Blauensteiner, M. Kampel, C. Musik, and S. Vogtenhuber, “A socio-technical approach for event detection in security critical infrastructure,” *2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. - Work. CVPRW 2010*, pp. 23–30, 2010.

[21] A. Laugé, J. Hernantes, and J. Mari Sarriegi, “The Role of Critical Infrastructures’ Interdependencies on the Impacts Caused by Natural Disasters,” *Crit. Inf. Infrastructures Secur.*, vol. 8328, pp. PP50–61, 2013.

[22] Novakovic. Jon, “The Impact Of Interdependence On Providing Protection For Critical Infrastructures,” 2015. .

[23] C. W. Johnson and K. McLean, “Tools for local critical infrastructure protection: computational support for identifying safety and security interdependencies between local critical infrastructures,” *3rd IET Int. Conf. Syst. Saf. 2008*, no. 2004, pp. 5A3–5A3, 2008.

[24] K. Smith, “Designing flexible curricula to enhance critical infrastructure security and resilience,” *Int. J. Crit. Infrastruct. Prot.*, vol. 7, no. 1, pp. 48–50, 2014.

[25] S. J. League, “Critical Infrastructure Protection - The Cyber/Information Dimension: Report on National Infrastructure Coordination Initiatives,” pp. 118–120, 1997.

[26] J. L. Rrushi, “An exploration of defensive deception in industrial communication networks,” *Int. J. Crit. Infrastruct. Prot.*, vol. 4, no. 2, pp. 66–75, 2011.

[27] L. Coppolino, S. D’Antonio, L. Romano, and G. Spagnuolo, “An Intrusion Detection System for Critical Information Infrastructures using Wireless Sensor Network technologies,” *Crit. Infrastruct. (CRIS), 2010 5th Int. Conf.*, pp. 1–8, 2010.

[28] B. C. Butler-Purry, S. Nuthalapati, and D. Kundur, “Network delay caused by cyber attacks on SVC and its impact on transient stability of smart grids,” *IEEE Power Energy Soc. Gen. Meet.*, vol. 2014-Octob, no. October, 2014.

[29] M. Kaâniche, “Resilience Assessment of Critical Infrastructures: From Accidental to Malicious Threats,” *CEUR Workshop Proc.*, vol. 1431, pp. 9–10, 2011.

[30] A. Tovey, “The Telegraph Industry Defence : Cyber attacks cost British industry £34bn a year,” 2015. .

[31] G. Giannopoulos, R. Filippini, and M. Schimmer, “Risk assessment methodologies for Critical Infrastructure Protection . Part I : A state of the art,” *JRC (Joint Res. Center) Eur. Comm. House Sci. Serv.*, p. 53, 2012.

[32] C. Copeland and B. Cody, “Terrorism and Security Issues Facing the Water Infrastructure Sector,” *Congr. Res. Serv. Rep.*, pp. 1–6, 2010.

[33] Environmenmental Protection Agency, “Planning for sustainability: A handbook for water and wastewater utilities,” no. February, pp. 1–75, 2012.

[34] P. Katsumata, J. Hemenway, W. Gavins, and B. Hamilton, “Cybersecurity risk management,” *Proc. - IEEE Mil. Commun. Conf. MILCOM*, pp. 890–895, 2010.

[35] J. C. Knight, “Safety critical systems: challenges and directions,” *Proc. 24th Int. Conf. Softw. Eng. (ICSE, 2002)*, pp. 547–550, 2002.

[36] E. E. L. Ii, J. E. Mitchell, and W. A. Wallace, “Restoration of Services in Interdependent Infrastructure Systems : A Network Flows Approach,” vol. 37, no. 6, pp. 1303–1317, 2007.

[37] L. Buttyan, D. Gessner, A. Hessler, and P. Langendoerfer, “Application of wireless sensor networks in critical infrastructure protection: challenges and design options [Security and Privacy in Emerging Wireless Networks,” *IEEE Wirel. Commun.*, vol. 17, no. 5, pp. 44–49, 2010.

[38] J. Xian, F. Lang, and X. Tang, “A NOVEL INTRUSION DETECTION METHOD BASED ON CLONAL SELECTION CLUSTERING ALGORITHM,” in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, 2005, no. August, pp. 18–21.

[39] M. Rong, C. Han, and L. Liu, “Critical infrastructure failure interdependencies in the 2008 Chinese Winter Storms,” *2010 Int. Conf. Manag. Serv. Sci. MASS 2010*, no. 70871093, p. 4, 2010.

[40] G. O’Reilly, A. Jrad, R. Nagarajan, T. Brown, and S. Conrad, “Critical infrastructure analysis of telecom for natural disasters,” *Networks 2006 12th Int. Telecommun. Netw. Strateg. Plan. Symp.*, pp. 1–6, 2007.

[41] S. Badri, P. Fergus, and W. Hurst, “A Cyber-Support System for Distributed Infrastructures,” in *EMERGING 2016 : The Eighth International Conference on Emerging Networks and Systems Intelligence*, 2016, p. 6.

[42] S. K. Badri, D. Llewellyn-jones, O. Abuelmaatti, and M. Merabti, “An Artificial Immune System Techniques for Critical Infrastructure Support,” in *The 14th Annual PostGraduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting (PGNet)*, 2013, p. 6.

[43] S. Badri, P. Fergus, and W. Hurst, “Statistical Analysis Methods for Interdependency Communication in Distributed Systems,” in *International Conference on Developments in eSystems Engineering*, 2016, p. 6.

[44] R. Zimmerman and L. Street, “Decision-making and the Vulnerability of Interdependent Critical ~ [nfrastructure ’,” pp. 4059–4063, 2004.

[45] S. Wang, L. Hong, X. Chen, J. Zhang, and Y. Yan, “Review of interdependent infrastructure systems vulnerability analysis,” *2011 2nd Int. Conf. Intell. Control Inf. Process.*, pp. 446–451, Jul. 2011.

[46] H. Teymourlouei, “Quick Reference : Cyber Attacks Awareness and Prevention Method for Home Users,” vol. 9, no. 3, pp. 480–486, 2015.

[47] A. Ramdin and A. Blackwell, “Report on Cybersecurity and Critical Infrastructure in the Americas,” *Trend Micro Incorporated*, p. 60, 2015.

[48] J. Dehlinger, J. B. Dugan, M. Veeraraghavan, M. McGinley, and C. L. Brown, "Continuous open design of dependable systems for critical infrastructure," *Proc. 2009 ICSE Work. Emerg. Trends Free. Source Softw. Res. Dev. FLOSS 2009*, pp. 48–53, 2009.

[49] J. Moteff, P. Parfomak, and I. Ave, "CRS Report for Congress Received through the CRS Web Critical Infrastructure and Key Assets : Definition and Identification," 2004.

[50] W. Hurst, M. Merabti, and P. Fergus, "Towards a Framework for Operational Support in Critical Infrastructures," *12th Annu. Postgrad. Symp. Converg. Telecommun. Netw. Broadcast. (PGNet 2011)*, 2011.

[51] T. Morris, A. Srivastava, B. Reaves, W. Gao, K. Pavurapu, and R. Reddi, "A control system testbed to validate critical infrastructure protection concepts," *Int. J. Crit. Infrastruct. Prot.*, vol. 4, no. 2, pp. 88–103, 2011.

[52] I. N. Fovino, M. Masera, L. Guidi, and G. Carpi, "An experimental platform for assessing SCADA vulnerabilities and countermeasures in power plants," *3rd Int. Conf. Hum. Syst. Interact. HSI'2010 - Conf. Proc.*, pp. 679–686, 2010.

[53] M. Polycarpou, G. Ellinas, E. Kyriakides, and C. Panayiotou, "Intelligent Health Monitoring of Critical Infrastructure Systems," *2010 Complex. Eng. Compeng 2010, Proc.*, pp. 18–20, 2010.

[54] D. Prochazkova, "Safety culture and critical infrastructure safety," *Proc. 2011 IEEE Int. Conf. Veh. Electron. Safety, ICVES 2011*, pp. 263–268, 2011.

[55] C. Scarlat, C. Simion, and E. I. Scarlat, "Managing new technology projects: Some considerations on risk assessment in the case of NPP critical infrastructures," *2011 2nd IEEE Int. Conf. Emerg. Manag. Manag. Sci.*, no. 836, pp. 911–915, 2011.

[56] O. Pavlovíc and H. D. Ehrich, "Model checking PLC software written in function block diagram," *ICST 2010 - 3rd Int. Conf. Softw. Testing, Verif. Valid.*, pp. 439–448, 2010.

[57] S. Zeng and Z. Yang, "A high performance architecture design of PLC dedicated processor," *ICACTE 2010 - 2010 3rd Int. Conf. Adv. Comput. Theory Eng. Proc.*, vol. 2, pp. 424–428, 2010.

[58] M. P. Coutinho, G. Lambert-Torres, L. E. Borges da Silva, J. G. Borges da Silva, J. C. Neto, and H. Lazarek, "Improving a methodology to extract rules to identify attacks in power system critical infrastructure: New results," in *2008 IEEE/PES Transmission and Distribution Conference and Exposition*, 2008, pp. 1–6.

[59] A. T. Murray and T. H. Grubesic, "Critical infrastructure protection: The vulnerability conundrum," *Telemat. Informatics*, vol. 29, no. 1, pp. 56–65, 2011.

[60] N. Lambert and K. Lin Song, "Use of query tokenization to detect and prevent SQL injection attacks," *Proc. - 2010 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2010*, vol. 2, pp. 438–440, 2010.

[61] H. Zhang, J. Ma, Y. Wang, and Q. Pei, "An active defense model and framework of insider threats detection and sense," *5th Int. Conf. Inf. Assur. Secur. IAS 2009*, vol. 1, pp. 258–261, 2009.

[62] European Network and Information Security Agency, "Cyber Europe 2010 – Evaluation Report Acknowledgments Contact details," *Inf. Secur.*, 2011.

[63] R. Sekar, T. Bowen, and M. Segal, "On preventing intrusions by process behavior monitoring," *Proc. Work. Intrusion Detect. Netw. Monit. (Id `99)*, pp. 29–40, 1999.

[64] L. Pla Beltran, M. Merabti, and W. Hurst, "Using Behavioural Observation and Game Technology to Support Critical Infrastructure Security," *Inderscience Int. J. Syst. Syst.*, vol. in press, no. 1, pp. 45–67, 2013.

[65] D. Dudenhoeffer, S. Hartley, and M. Permann, "Critical Infrastructure Interdependency Modeling : A Survey of Critical Infrastructure Interdependency Modeling : A," no. August, 2006.

[66] European Commission, "Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions: For a European Industrial Renaissance," p. 25, 2014.

[67] A. Di Giorgio and F. Liberati, "Interdependency modeling and analysis of critical infrastructures based on Dynamic Bayesian Networks," in *2011 19th Mediterranean Conference on Control & Automation (MED)*, 2011, pp. 791–797.

[68] C. Han, L. Liu, and M. Rong, "Addressing Criticality Levels in Critical Infrastructure System," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, no. October, pp. 3965 – 3970.

[69] M. Dogrul, A. Aslan, and E. Celik, "Developing an international cooperation on cyber defense and deterrence against Cyber terrorism," *2011 3rd Int. Conf. Cyber Confl.*, pp. 1–15, 2011.

[70] N. Abouzakhar, "Critical Infrastructure Cybersecurity: A Review of Recent Threats and Violations," 2015, vol. 1, p. 11.

[71] S. J. Yang, J. Holsopple, and M. Sudit, "Evaluating threat assessment for multi-stage cyber attacks," *Proc. - IEEE Mil. Commun. Conf. MILCOM*, 2006.

[72] B. Chen, K. L. Butler-Purry, and D. Kundur, "Impact analysis of transient stability due to cyber attack on FACTS devices," *2013 North Am. Power Symp.*, pp. 1–6, 2013.

[73] J. Smith, J. Pereyda, and D. Gammel, “Cybersecurity best practices for creating resilient control systems,” *Proc. - 2016 Resil. Week, RWS 2016*, pp. 62–66, 2016.

[74] Pwc, “Moving forward with cybersecurity and privacy,” 2016.

[75] D. Wanger, “The Growing Threat of Cyber-Attacks on Critical Infrastructure,” 2016. .

[76] M. Feily, A. Shahrestani, and S. Ramadass, “A survey of botnet and botnet detection,” *Proc. - 2009 3rd Int. Conf. Emerg. Secur. Information, Syst. Technol. Secur. 2009*, pp. 268–273, 2009.

[77] C. Steitz and E. Auchard, “German nuclear plant infected with computer viruses, operator says,” 2016. .

[78] Price Waterhouse Coopers, “Cyber attacks Is your critical infrastructure safe?,” pp. 1–7, 2010.

[79] M. Van Doorn, “Resilient wireless data communication for critical infrastructure,” *Isgt 2011*, pp. 1–5, 2011.

[80] X. Li, X. Liang, R. Lu, X. Shen, X. Lin, and H. Zhu, “Securing smart grid: Cyber attacks, countermeasures, and challenges,” *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, 2012.

[81] H. Chen, R. Chiang, and V. C. Storey, “BUSINESS INTELLIGENCE AND ANALYTICS : FROM BIG DATA TO BIG IMPACT,” *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2012.

[82] S. Jagannathan, “Real-Time Big Data Analytics Architecture For Remote Sensing Application,” *Int. Conf. Signal Process. Commun. Power Embed. Syst.*, pp. 1912–1916, 2016.

[83] A. Tole, “Big Data challenges,” *Database Syst.*, vol. 4, no. 3, pp. 31–40, 2013.

[84] H. Mishra, “Big Data: Issues and Challenges,” 2014. .

[85] M. Hibsh, “Overview of the huge data,” *the world of Technology*, 2013. .

[86] W. Swapnil, Y. Anil, and S. Gupta, “Big Data: Characteristics, Challenges and Data Mining,” *Int. J. Comput. Appl.*, no. September, pp. 25–29, 2016.

[87] M. Rouse, “big data analytics,” *SearchBusinessAnalytics*, 2017. .

[88] British Academy for Training and Development, “big data: analytics,” 2017. .

[89] SAS Post, “How the analysis of the giant data is a breakthrough in the performance of companies and jobs,” 2015. .

[90] S. Pyne, B. L. . Rao, and S. . Rao, *Big Data Analytics: Methods and Applications*, Illustrate. Springer, 2016.

[91] EMC Education Services, *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*, Illustrate. John Wiley & Sons, 2015.

[92] T. Maydon, "The Data Analytics Blog," *principa a transaction capital company*, 2017. .

[93] S. Bangsow, *Manufacturing Simulation with Plant Simulation and SimTalk*, vol. 1. 2015.

[94] S. D. Wolthusen, "Modeling Critical Infrastructure Requirements," no. June, pp. 10–11, 2004.

[95] D. Asteteh and O. Sarhan, *Education and e-Learning Technology*. Jorden: darwael, 2007.

[96] A. AL-esaa, "The effect of using simulation Implementing strategy through computer teaching assistant in the immediate and delayed achievement," Jorden, 1993.

[97] M. Hall-May and M. Surridge, "Resilient Critical Infrastructure Management Using Service Oriented Architecture," *2010 Int. Conf. Complex, Intell. Softw. Intensive Syst.*, pp. 1014–1021, 2010.

[98] Siemens, "www.siemens.com/tecnomatix," 2011.

[99] C. M. Lewandowski, N. Co-investigator, and C. M. Lewandowski, *Manufacturing Simulation with Plant Simulation and SimTalk*, vol. 1. 2015.

[100] Industrial Automation Systems, "Modicon modbus protocol reference guide," *Industrial Automation Systems*, 1996.

[101] W. Hurst, M. Merabti, and P. Fergus, "Behavioural Observation for Critical Infrastructure Support.," in *13th Annual Postgraduate Symposium on the Convergence of Tele-communications, Networking and Broadcasting*, 2012.

[102] A. Kordon, *Applying Computational Intelligence: How to Create Value*. USA: Springer, 2010.

[103] J. Hernantes, A. Laugé, L. Labaka, E. Rich, F. O. Sveen, J. Mari, I. J. Martinez-moyano, and J. J. Gonzalez, "Collaborative Modeling of A wareness in Critical Infrastructure Protection University of Chicago," pp. 1–10, 2011.

[104] R. Klein, E. Rome, and C. Beyel, "Information Modelling and Simulation in large interdependent Critical Infrastructures in IRRIIS," pp. 1–21, 2009.

[105] M. Chaturvedi, "Cyber Security Infrastructure in India: A Study," *Emerg. Technol. ...*, pp. 70–84, 2008.

[106] J. P. Marchal, Gu., Marchel, G. and Heiken, *Multidetector-Row Computed Tomography: Scanning and Contrast Protocols*. Italia: Springer-Verlag, 2005.

[107] L. Sompayrac, *How the Immune System Works*, 4th ed. John Witey and Sons, 2012.

[108] E. B. Fernandez and M. M. Larrondo-Petrie, "Designing secure SCADA systems using security patterns," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, pp. 1–8, 2010.

[109] M. Elsadig and A. Abduallah, "Biological Intrusion Prevention and Self-Healing Model for Network Security," in *2010 Second International Conference on Future Networks*, 2010, pp. 337–342.

[110] E. Commission, "Digital Agenda: cyber-security experts test defences in first pan-European simulation," *Eur. Comm. Press Release*, no. November, 2010.

[111] C. M. Davis, J. E. Tate, H. Okhravi, C. Grier, T. J. Overbye, and D. Nicol, "SCADA cyber security testbed development," in *2006 38th Annual North American Power Symposium, NAPS-2006 Proceedings*, 2006, pp. 483–488.

[112] Ç. ÜNDEĞER, "Introduction To Modeling & Simulation," in *Bilkent Üniversitesi*, 2008, no. Part 1, pp. 1–51.

[113] National Institute of Standards and Technology, "Framework for improving critical infrastructure cybersecurity," no. April, 2015.

[114] J.-C. Laprie, K. Kanoun, and M. Kaaniche, "Modelling interdependencies between the electricity and information infrastructures," *26th Int. Conf. Comput. Safety, Reliab. Secur.*, vol. abs/0809.4, pp. 1–14, 2008.

[115] I. Dobson, B. A. Carreras, and D. E. Newman, "a Loading-Dependent Model of Probabilistic Cascading Failure," *Probab. Eng. Informational Sci.*, vol. 19, no. 01, pp. 475–488, 2005.

[116] I. Dobson and B. A. Carreras, "Risk analysis of critical loading and blackouts with cascading events(CASCADE)," 2005.

[117] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos*, vol. 12, no. 4, pp. 985–994, 2002.

[118] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, no. 4, pp. 1–4, 2004.

[119] R. Kinney, P. Crucitti, R. Albert, and V. Latora, "Modeling cascading failures in the North American power grid," *Eur. Phys. J. B*, vol. 46, no. 1, pp. 101–107, 2005.

[120] D. P. Chassin and C. Posse, “Evaluating North American electric grid reliability using the Barabasi-Albert network model,” *Physica A*, vol. 355, no. 2–4, pp. 667–677, 2005.

[121] A. Müller and S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O’Reilly Media, Inc., 2017.

[122] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, “Supervised Machine Learning : A Review of Classification Techniques,” *Emerg. Artif. Intell. Appl. Comput. Eng.*, vol. 31, pp. 3–4, 2007.

[123] Z. Xu, I. King, S. Member, and M. R. Lyu, “Discriminative Semi-Supervised Feature Selection Via Manifold Regularization,” vol. 21, no. 7, pp. 1033–1047, 2010.

[124] S. Raschka, *Python Machine Learning*. Packt Publishing Ltd, 2015.

[125] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, “Supervised machine learning: A review of classification techniques,” vol. 31, pp. 249–268, 2007.

[126] A. Singh, N. Thakur, and A. Sharma, “A review of supervised machine learning algorithms,” *Proc. 10th INDIACom; 2016 3rd Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2016*, pp. 1310–1315, 2016.

[127] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, “Machine learning: a review of classification and combining techniques,” *Artif. Intell. Rev.*, vol. 26, no. 3, pp. 159–190, Nov. 2007.

[128] Y. Y. Haimes, J. R. Santos, K. G. Crowther, and M. H. Henry, “Analysis of Interdependencies and Risk in Oil & Gas Infrastructure Systems,” Virginia, 2007.

[129] Y. Y. Haimes, J. Lambert, D. Li, R. Schooff, and V. Tulsani, “Hierarchical holographic modeling for risk identification incomplex systems,” *1995 IEEE Int. Conf. Syst. Man Cybern.*, vol. 2, pp. 606–617, 1981.

[130] C. T. Schauppner, “Measuring the Immeasurable : Applying Hierarchical Holographic Modeling to Developing Measures of Effectiveness for Stability , Security , Transition , and Reconstruction Operations,” Newport, 2006.

[131] E. J. Byres, M. Franz, and D. Miller, “The use of attack trees in assessing vulnerabilities in SCADA systems,” in *International Infrastructure Survivability Workshop (IISW’04), Institute for Electrical and Electronics Engineers*, 2004, pp. 1–9.

[132] S. Badri, P. Fergus, and W. Hurst, “Critical Infrastructure Automated Immuno-Response System (CIAIRS),” in *3rd International conference on control, Decision and Information Technologies (CoDIT), Malta*, 2016, p. 6.

[133] A. R. McGee, S. Rao Vasireddy, C. Xie, D. D. Picklesimer, U. Chandrashekhar, and S. H. Richman, “A framework for ensuring network security,” *Bell Labs Tech. J.*, vol. 8, no. 4, pp. 7–27, 2004.

[134] E. D. Knapp and J. Langill, *Industrial Network Security: Securing Critical Infrastructure Networks for Smart Grid, SCADA, and Other Industrial Control Systems*, 2, illustr. Elsevier Science Limited, 2014, 2014.

[135] M. Al-Ahwal, *Infrastructure failures of water networks in Saudi Arabia*. 2014.

[136] S. Badri, P. Fergus, and W. Hurst, “A Support Network for Distributed Systems,” *10th Int. Conf. E-Learning Games,Edutainment 2016.*, p. 16, 2016.

[137] The USGS Water Science School, “Hydroelectric power: How it works,” *U.S. Departement of the Interior*, 2016. .

[138] S.moazzem, R. M.G., and K. M.M, “A Review on Technologies for Reducing CO2 Emission from Coal Fired Power Plants,” in *Thermal Power Plants*, M. Rasul, Ed. 2012.

[139] 45 Nuclear Plants, “Nuclear Reactor Designs,” *Nuclear the new Green Energy*, 2017. .

[140] J. Lin, S. Sedigh, and A. Miller, “Integrated Cyber-Physical Simulation of Intelligent Water Distribution Networks,” *Proc. 2009 Eighth IEEE Int. Conf. Dependable, Auton. Secur. Comput.*, pp. 690–695, 2009.

[141] L. Phoenix and W. Atkins, “Infrastructure, Water-Supply,” *Water Encyclopedia Science and Issues*, 2017. .

[142] N. K. Svendsen and S. D. Wolthusen, “Analysis and statistical properties of critical infrastructure interdependency multiflow models,” *Proc. 2007 IEEE Work. Inf. Assur. IAW*, no. June, pp. 247–254, 2007.

[143] J. Pallant, *Spss Survival Manual*, Illustrate. USA: McGraw-Hill Education (UK), 2013.

[144] P. Fergus, P. Cheung, P. Hussain, D. Al-Jumeily, C. Dobbins, and S. Iram, “Prediction of Preterm Deliveries from EHG Signals Using Machine Learning,” *PLoS One*, vol. 8, no. 10, p. e77154, 2013.

[145] X. Wang and K. K. Paliwal, “Feature extraction and dimensionality reduction algorithms and their applications in vowel recognition,” *Pattern Recognit.*, vol. 36, no. 10, pp. 2429–2439, 2003.

[146] F. Lotte, “Study of Electroencephalographic Signal Processing and Classification Techniques towards the use of Brain-Computer Interfaces in Virtual Reality Applications,” 2008.

APPENDIX A-DATA

To provide a bigger view regarding the Data, a large sample of the normal and abnormal datasets are presented for two Critical Infrastructures, the Hydroelectricity System and the Electricity System, as an example. In addition, the faults presented for both systems.

Tecnomatix Simulator Data

Hydroelectricity System

Table 26: The Hydroelectricity System Abnormal Behaviour

Failure	Components	Available	Not Available
1	Water pipe 1	40	60
	Turbine	50	50

Table 27: Component Descriptive for The Hydroelectricity System

Abbreviation	Components Description
F1	water pipe from the Swage system to the Hydroelectricity system in the Hydroelectricity
F2	buffer in the Hydroelectricity system
F3	water pipe 1 from sewage in the Hydroelectricity system
F4	sluice-gate in the hydroelectricity system
F5	water pipe 2 in the Hydroelectricity
F6	the turbine in the Hydroelectricity system

F7	hydroelectricity generator
F8	hydroelectricity transformer
F9	electricity cable in the hydroelectricity system

Table 28: Large Data Sample for Normal Hydroelectricity System

Time	F1	F2	F3	F4	F5	F6	F7	F8	F9
42:33.5	0	21	3	40	3	1	1	1	0
42:33.8	0	21	3	40	3	1	1	1	0
42:34.0	0	21	3	40	3	1	1	1	0
42:34.2	0	21	3	40	3	1	1	1	0
42:34.5	0	21	3	40	3	1	1	1	0
42:34.7	0	21	3	40	3	1	1	1	0
42:35.0	0	21	3	40	3	1	1	1	0
42:35.3	0	21	3	40	3	1	1	1	0
42:35.5	0	21	3	40	3	1	1	1	0
42:35.8	0	21	3	40	3	1	1	1	0
42:36.0	0	21	3	40	3	1	1	1	0
42:36.3	0	21	3	40	3	1	1	1	0
42:36.5	0	21	3	40	3	1	1	1	0
42:36.7	0	21	3	40	3	1	1	1	0
42:37.0	0	21	3	40	3	1	1	1	0
42:37.3	0	21	3	40	3	1	1	1	0
42:37.5	0	21	3	40	3	1	1	1	0
42:37.7	0	21	3	40	3	1	1	1	0
42:38.0	0	21	3	40	3	1	1	1	0
42:38.3	0	21	3	40	3	1	1	1	0
42:38.5	0	21	3	40	3	1	1	1	0
42:38.8	0	21	3	40	3	1	1	1	0
42:39.0	0	21	3	40	3	1	1	1	0
42:39.3	0	21	3	40	3	1	1	1	0
42:39.5	0	21	3	40	3	1	1	1	0
42:39.7	0	21	3	40	3	1	1	1	0
42:40.0	0	21	3	40	3	1	1	1	0

42:40.3	0	21	3	40	3	1	1	1	0
42:40.5	0	21	3	40	3	1	1	1	0
42:40.7	0	21	3	40	3	1	1	1	0
42:41.0	0	21	3	40	3	1	1	1	0
42:41.2	0	21	3	40	3	1	1	1	0

Table 29: Large Data Sample for Abnormal Hydroelectricity System

Time	F1	F2	F3	F4	F5	F6	F7	F8	F9
42:33.5	3	40	3	40	3	1	0	1	1
42:33.8	3	40	3	40	3	0	1	1	1
42:34.0	3	40	3	40	3	0	1	1	1
42:34.2	3	40	3	40	3	0	1	1	1
42:34.5	3	40	3	40	3	0	1	1	1
42:34.7	3	40	3	40	3	0	1	1	1
42:35.0	3	40	3	40	3	0	1	1	1
42:35.3	3	40	3	40	3	0	1	1	1
42:35.5	3	40	3	40	3	0	1	1	1
42:35.8	3	40	3	40	2	1	1	1	0
42:36.0	3	40	2	40	3	1	1	1	0
42:36.3	3	40	2	40	3	1	1	1	0
42:36.5	2	40	3	40	3	1	1	1	0
42:36.7	2	40	3	40	3	1	1	1	0
42:37.0	2	40	3	40	3	1	1	1	0
42:37.3	3	40	3	40	3	1	1	1	0
42:37.5	3	40	3	40	3	1	1	1	0
42:37.7	3	40	3	40	3	1	1	1	0
42:38.0	3	40	3	40	3	1	1	1	0
42:38.3	3	40	3	40	3	1	1	1	0
42:38.5	3	40	3	40	3	1	1	1	0
42:38.8	3	40	3	40	3	1	1	1	0
42:39.0	3	40	3	40	3	1	1	1	0
42:39.3	3	40	3	40	3	1	1	1	0
42:39.5	3	40	3	40	3	1	1	1	0
42:39.7	3	40	3	40	3	1	1	1	0

42:40.0	3	40	3	40	3	1	1	1	0
42:40.3	3	40	3	40	3	1	1	1	0
42:40.5	3	40	3	40	3	1	1	1	0
42:40.7	3	40	3	40	3	1	1	1	0
42:41.0	3	40	3	40	3	1	1	1	0
42:41.2	3	40	3	40	3	1	1	1	0

Electricity System

Table 30: The Electricity System Abnormal Behaviour

Failure	Components	Available	Not Available
1	Electricity cable from nuclear power	70	30
	Electricity cable to water distribution	60	40
	Electricity cable to factory	80	20

Table 31: Component Descriptive for The Electricity System

Abbreviation	Components Descriptions
F1	electricity cable from the hydroelectricity to the SmartGrid in the Grid
F2	electricity cable from the Nuclear to the SmartGrid in the Grid
F3	electricity cable from the Coal to the SmartGrid in the Grid
F4	the buffer in the SmartGrid
F5	electricity cable from the SmartGrid to the Factory in the Grid
F6	electricity cable from the SmartGrid to the Houses in the Grid
F7	electricity cable from the SmartGrid to the WD in the Grid

Table 32: Large Data Sample for Normal Electricity System

Time	F1	F2	F3	F4	F5	F6	F7
1:22:48:12.7500	0	0	0	0	0	0	3
1:22:48:13.0000	0	0	0	0	0	0	3
1:22:48:13.2500	0	0	0	0	0	0	3
1:22:48:13.5000	0	0	0	0	0	0	3
1:22:48:13.7500	0	0	0	0	0	0	3
1:22:48:14.0000	0	0	0	0	0	0	2
1:22:48:14.2500	0	0	0	0	0	0	2
1:22:48:14.5000	0	0	0	0	0	0	2
1:22:48:14.7500	0	0	0	0	0	0	2
1:22:48:15.0000	0	0	0	0	0	0	1
1:22:48:15.2500	0	0	0	0	0	0	1
1:22:48:15.5000	0	0	0	0	0	0	1
1:22:48:15.7500	0	0	0	0	0	0	1
1:22:48:16.0000	0	0	0	0	0	0	1
1:22:48:16.2500	0	0	0	0	0	0	1
1:22:48:16.5000	0	0	0	0	0	0	1
1:22:48:16.7500	0	0	0	0	0	0	1
1:22:48:17.0000	0	0	0	0	0	0	1
1:22:48:17.2500	0	0	0	0	0	0	1
1:22:48:17.5000	0	0	0	0	0	0	1
1:22:48:17.7500	0	0	0	0	0	0	1
1:22:48:18.0000	0	0	0	0	0	0	1
1:22:48:18.2500	0	0	0	0	0	0	1
1:22:48:18.5000	0	0	0	0	0	0	1
1:22:48:18.7500	0	0	0	0	0	0	1
1:22:48:19.0000	0	0	0	0	0	0	1
1:22:48:19.2500	0	0	0	0	0	0	1
1:22:48:19.5000	1	0	0	0	0	0	1
1:22:48:19.7500	1	0	0	0	0	0	1
1:22:48:20.0000	1	0	0	0	0	0	1
1:22:48:20.2500	1	0	0	0	0	0	1
1:22:48:20.5000	1	0	0	0	0	0	1

1:22:48:20.7500	1	0	0	0	0	0	1
1:22:48:21.0000	0	0	0	1	0	0	1
1:22:48:21.2500	0	0	0	1	0	0	1
1:22:48:21.5000	0	0	0	1	0	0	1
1:22:48:21.7500	0	0	0	1	0	0	1

Table 33: Large Data Sample for Abnormal Electricity System

Time	F1	F2	F3	F4	F5	F6	F7
1:22:48:12.7500	1	0	0	0	0	0	3
1:22:48:13.0000	1	0	0	0	0	0	3
1:22:48:13.2500	1	0	0	0	0	0	3
1:22:48:13.5000	1	0	0	0	0	0	3
1:22:48:13.7500	1	0	0	0	0	0	3
1:22:48:14.0000	1	0	0	0	0	0	3
1:22:48:14.2500	0	0	0	1	0	0	3
1:22:48:14.5000	0	0	0	1	0	0	3
1:22:48:14.7500	0	0	0	1	0	0	3
1:22:48:15.0000	0	0	0	1	0	0	3
1:22:48:15.2500	0	0	0	0	1	0	3
1:22:48:15.5000	0	0	0	0	1	0	3
1:22:48:15.7500	0	0	0	0	1	0	3
1:22:48:16.0000	0	0	0	0	1	0	3
1:22:48:16.2500	0	0	0	0	1	0	3
1:22:48:16.5000	0	0	0	0	1	0	3
1:22:48:16.7500	0	0	0	0	1	0	3
1:22:48:17.0000	0	0	0	0	1	0	3
1:22:48:17.2500	0	0	0	0	1	0	3
1:22:48:17.5000	0	0	0	0	1	0	3
1:22:48:17.7500	0	0	0	0	0	0	3
1:22:48:18.0000	0	0	0	0	0	0	3
1:22:48:18.2500	0	0	0	0	0	0	3
1:22:48:18.5000	0	0	0	0	0	0	2
1:22:48:18.7500	0	0	0	0	0	0	2
1:22:48:19.0000	0	0	0	0	0	0	2

1:22:48:19.2500	0	0	0	0	0	0	2
1:22:48:19.5000	0	0	0	0	0	0	1
1:22:48:19.7500	0	0	0	0	0	0	1
1:22:48:20.0000	0	0	0	0	0	0	1
1:22:48:20.2500	0	0	0	0	0	0	1
1:22:48:20.5000	0	0	0	0	0	0	1
1:22:48:20.7500	0	0	0	0	0	0	1
1:22:48:21.0000	0	0	0	0	0	0	1
1:22:48:21.2500	0	0	0	0	0	0	1
1:22:48:21.5000	0	0	0	0	0	0	1
1:22:48:21.7500	0	0	0	0	0	0	1

APPENDIX B-FIGURES

This section presents the Simulation Plant for the Hydroelectricity System and the Electricity System. Figure 67 and 68 demonstrates the abnormal flow in Hydroelectricity and the Electricity systems, respectively, depending on the abnormal behaviour introduced to the components, which are presented in Table 29 & 33. A Bottleneck Analyser presents a statistical data graphically on top of each component.

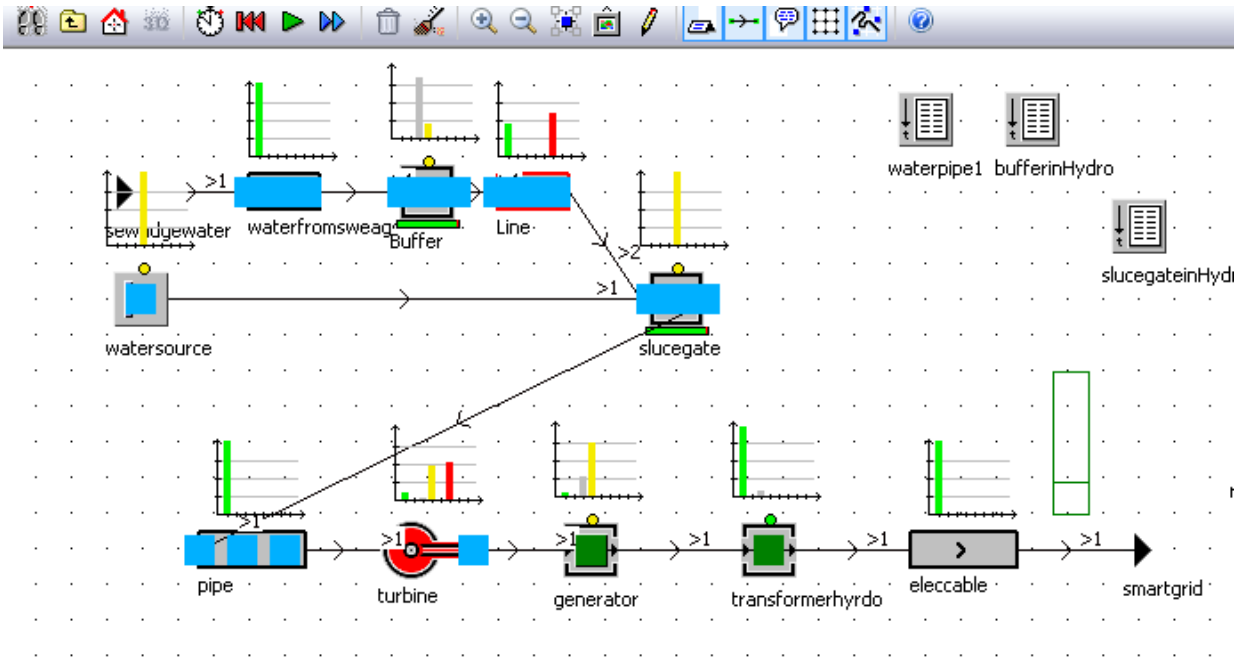


Figure 67 The abnormal behaviour in the Hydroelectricity System: Water pipe 1 & Turbine.

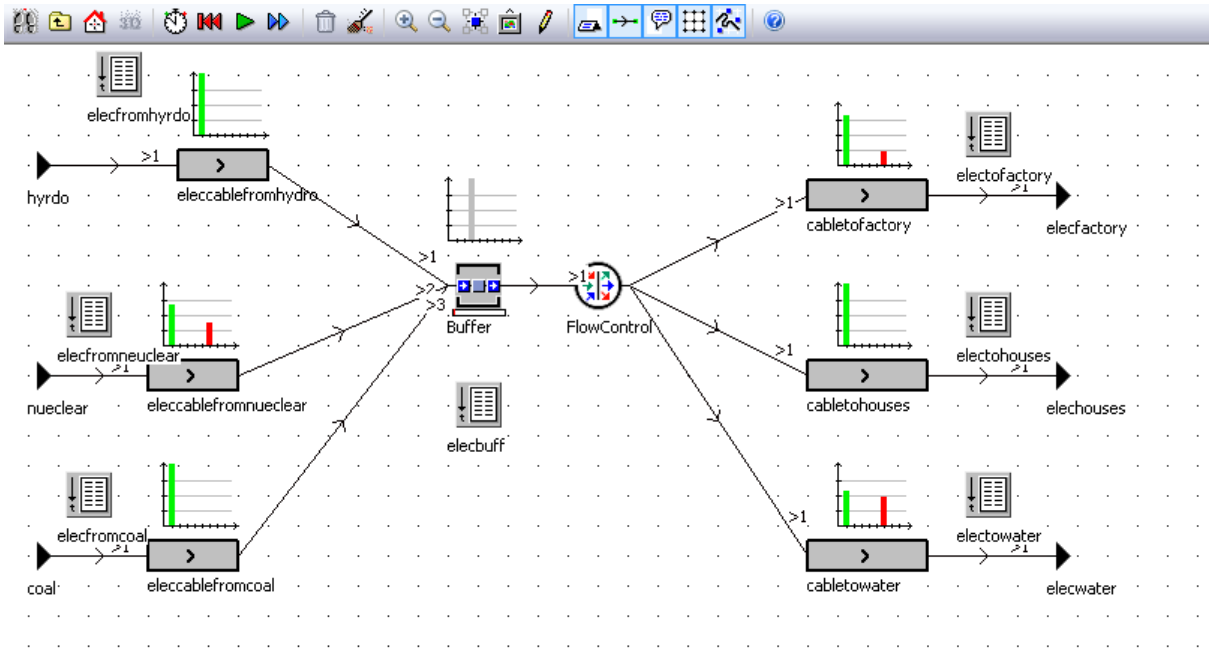


Figure 68 The Abnormal Behaviour in the Electricity System: electricity cable from the Nuclear system, electricity cable to the Factory and the Houses

Figure 69 and 70 indicates the abnormal behaviour for the Hydroelectricity and the Electricity systems, respectively, effect on the rest of the components.

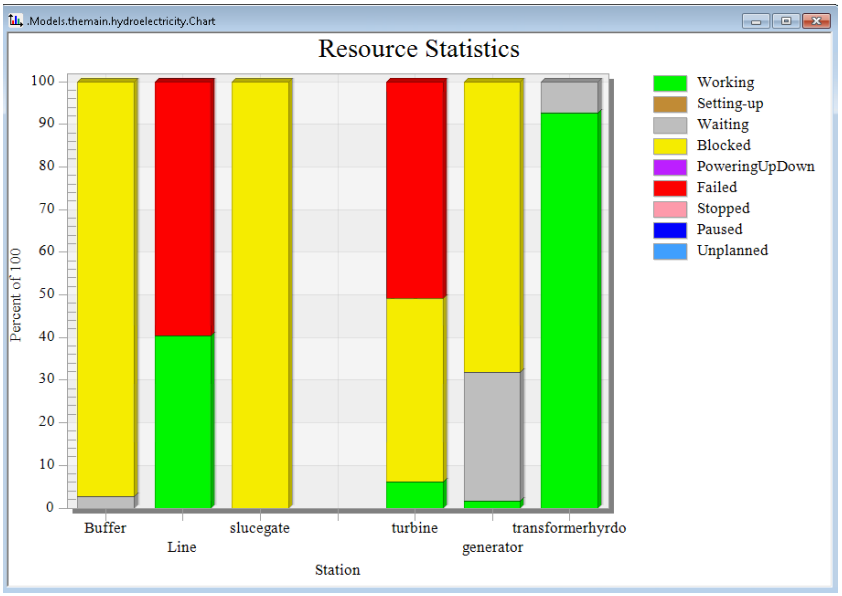


Figure 69 The Abnormal Chart for the Hydroelectricity System: Water pipe 1 & Turbine

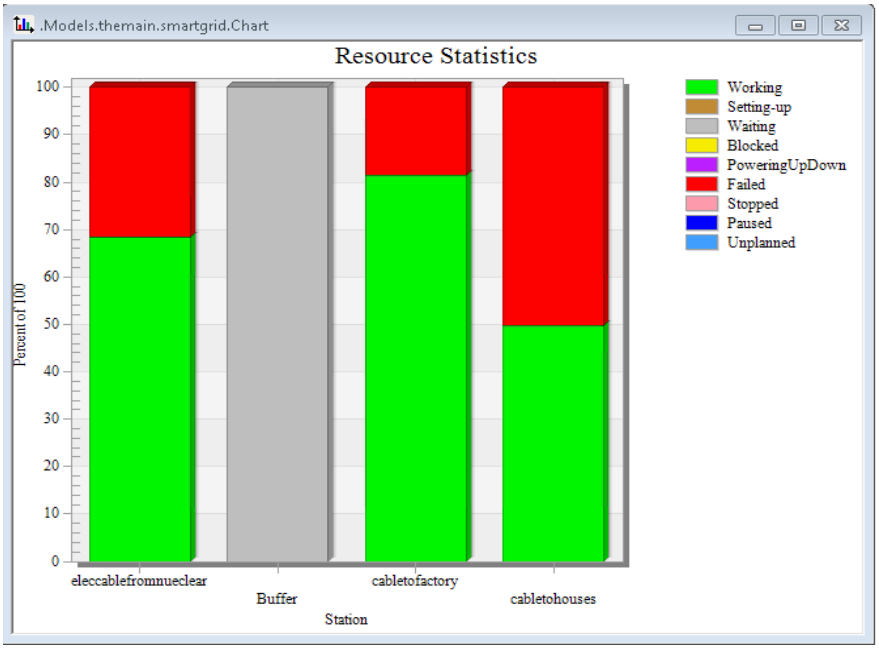


Figure 70 The Abnormal Chart for the Electricity System

APPENDIX C-RECORDS

This section presents the full 864 vector records for both the normal and abnormal water distribution system.

Table 34 The Normal Water Distribution System Vector Records

normal	electricity	std	water pipe	std	water pipe	std	WD treatn	std	water pipe	std	WD Storag	std	water pipe	std	water pipe	std	water pipe	std
Record 1	0.053333	0.224791	2.975	0.185482	0.066667	0.249548	0.033333	0.17958	0.066667	0.249548	0.033333	0.17958	0.08	0.271406	0.033333	0.17958	0.033333	0.17958
Record 2	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.008326	0.090906
Record 3	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.011657	0.107381
Record 4	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.009992	0.099499
Record 5	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.009992	0.099499
Record 6	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.011657	0.107381
Record 7	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.008326	0.090906
Record 8	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 9	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 10	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 11	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 12	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 13	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 14	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 15	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 16	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 17	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 18	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 19	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 20	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 21	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 22	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 23	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 24	0.02086	0.142963	2.99416	0.114187	0.0267	0.161258	0.01335	0.114809	0.0267	0.161258	0.01335	0.114809	0.03254	0.177487	0.01335	0.114809	0.007512	0.086375

Table 35 The Abnormal Water Distribution System Vector Records

abnormal	electricity	std	water pipe	std	water pipe	std	WD treatn	std	water pipe	std	WD Storag	std	water pipe	std	water pipe	std	water pipe	std
Record 1	0.338333	0.681359	2.9775	0.178943	0.066667	0.286864	0.033333	0.17958	0.066667	0.286864	0.033333	0.17958	0.0775	0.316086	0.033333	0.17958	0.033333	0.17958
Record 2	0.020816	0.142827	2.995004	0.070534	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.008326	0.090906
Record 3	0.067444	0.250893	2.995004	0.070534	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.016653	0.12802	0.006661	0.081377	0.011657	0.107381
Record 4	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.009992	0.099499	0.013322	0.114698
Record 5	0.010824	0.103518	2.999167	0.028855	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.009992	0.099499	0.006661	0.081377
Record 6	0.174022	0.379286	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.03164	0.175113	0.006661	0.081377	0.018318	0.134155
Record 7	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.008326	0.090906
Record 8	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 9	0.010824	0.103518	2.769359	0.421419	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 10	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.006661	0.081377	0.019983	0.140001
Record 11	0.060783	0.239031	2.996669	0.057639	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 12	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 13	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 14	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 15	0.010824	0.103518	2.996669	0.057639	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 16	0.020816	0.142827	2.994172	0.076153	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 17	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 18	0.020816	0.142827	2.995004	0.070534	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 19	0.010824	0.103518	2.997502	0.049938	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 20	0.020816	0.142827	2.995004	0.070534	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.013322	0.114698	0.013322	0.114698
Record 21	0.010824	0.103518	2.996669	0.057639	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 22	0.020816	0.142827	2.993339	0.081377	0.026644	0.161109	0.013322	0.114698	0.026644	0.161109	0.013322	0.114698	0.032473	0.177326	0.006661	0.081377	0.019983	0.140001
Record 23	0.010824	0.103518	2.998335	0.040791	0.013322	0.114698	0.006661	0.081377	0.013322	0.114698	0.006661	0.081377	0.01582	0.124831	0.006661	0.081377	0.006661	0.081377
Record 24	0.02086	0.142963	2.99416	0.114384	0.0267	0.161258	0.01335	0.114809	0.0267	0.161258	0.01335	0.114809	0.03254	0.177487	0.01335	0.114809	0.007512	0.086375