DESIGN AND ANALYSIS OF SECURE DIGITAL TWIN ARCHITECTURE

by

Eman Shaikh

A Thesis presented to the Faculty of the
American University of Sharjah
College of Engineering
In Partial Fulfilment
of the Requirements
for the Degree of

Master of Science in
Computer Engineering

Sharjah, United Arab Emirates

June 2022

Declaration of Authorship

I declare that this thesis is my own work and, to the best of my knowledge and belief, it does not contain material published or written by a third party, except where permission has been obtained and/or appropriately cited through full and accurate referencing.

Eman Shaikh
Signed…………………………………………………………….

28/06/21
Date…………………………………………………………

Approval Signatures

We, the undersigned, approve the Master's Thesis of Eman Shaikh

Thesis Title: Design and Analysis of Secure Digital Twin Architecture

Date of Defence: 24/05/22

Name, Title and Affiliation	**Signature**
Dr. Abdulrahman Al-Ali Professor, Department of Computer Science and Engineering Thesis Advisor	
Dr. Nazeeruddin Mohammad Associate Professor, College of Computer Science and Engineering Prince Mohammad bin Fahd University Thesis Co-Advisor	
Dr. Salam Dhou Assistant Professor, Department of Computer Science and Engineering Thesis Committee Member	
Dr. Lutfi Albasha Professor, Department of Electrical Engineering Thesis Committee Member	
Dr. Imran Zualkernan Interim Department Head Department of Computer Science and Engineering	
Dr. Lotfi Romdhane Associate Dean for Graduate Affairs and Research College of Engineering	
Dr. Fadi Aloul Dean College of Engineering	
Dr. Mohamed El-Tarhuni Vice Provost for Research and Graduate Studies Office of Research and Graduate Studies	

Acknowledgements

First and foremost, I would like to praise Allah the Almighty, the Most Gracious and the Most Merciful for the blessings He has bestowed upon me throughout the completion of my graduate program.

Next, I would like to thank my parents, family, and friends for their never-ending support and encouragement throughout the years.

I would like to express my sincere gratitude to my advisors Dr. Abdulrahman Al-Ali and Dr. Nazeeruddin Mohammad for providing me their continuous support, knowledge, guidance, and encouragement throughout my research stages. I'm deeply beholden for having them as my advisors. I would also like to thank Dr. Shahabuddin Muhammad for providing me knowledge and assistance whenever I needed help.

I would like to extend my regards to Dr. Salam Dhou and Dr. Lutfi Albasha for serving as a committee member and providing their valuable feedback.

Lastly, I would like to thank the American University of Sharjah and the Department of Computer Science and Engineering for offering me the Graduate Teaching Assistantship during my master studies.

Abstract

In recent years, Digital Twin has been gaining popularity in various smart city applications like smart manufacturing, smart energy, smart mobility, and smart healthcare. Digital Twin in simple term, can be described as a virtual replica of a given physical product, system, or process. It comprises of three major components: the object entity, data acquisition unit and the virtual counterpart. The virtual entity is a cloud-based layer that performs data analytic and visualization. The data analytic utilizes Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) algorithms that help in predictive maintenance, optimization, and improved productivity. While the collected data from the physical domain layer is transmitted, exchanged, and processed at the virtual domain and application layers, it will be exposed to several security attacks. Nevertheless, just like any emerging technology, several research studies have been conducted over the implementation of Digital Twin in many smart city applications without paying enough attention in developing a model that comprises of the security attacks, analysis, and remedies. Therefore, the objective of this thesis is to propose a secure Digital Twin multi-layer architecture for the attacks that can occur in each layer. It also proposes and implements a Probabilistic Model Checking (PMC) based approach to assess the Digital Twin security. The proposed model will be used to analyse the probability of success, as well as the cost for a variety of potential attacks that can occur in the Digital Twin architecture. To assess the proposed solution, a case study on an application of Digital Twin in healthcare will be considered.

Keywords: Digital twins; security; cybersecurity; modelling; healthcare.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

EHR	Electronic Health Record
EMR	Electronic Medical Record
IoT	Internet of Things
PHR	Personal Health Record
WEP	Wired Equivalent Privacy

Chapter 1. Introduction

This chapter provides an overview on the concept of Digital Twin. Next, it presents the problem statement of the thesis. Then, it proceeds to define the objective of the thesis. Lastly, it outlines the organization followed in the thesis report.

1.1. Overview

In recent years, an exponential rise in the global population has led to the advancement towards various enabling technologies. CPS is one such technology whose role is to integrate the cyber world with the physical world [1]. This integration helps to provide real-time sensing, information feedback and dynamic control to the given physical entity. Since most of the products present in the market today comprises of sensors, actuators, micro-controller, along with internet connectivity. Industry 4.0 refers to such product as a smart product that consists of both a physical part and a virtual counterpart, that is now more commonly known as the Digital Twin (DT).

Furthermore, the advancement in the digital world has made connectivity as one of the most important features. This is because it helps to create a bond between the physical and the digital world. That is, through connectivity, human can now interact with technology like never before. DT and IoT are some of the two major enabling technology that features in the digital representation. The major difference between them is that unlike IoT, DT is considered as a virtual representation of any given physical entity in real-time which is obtained through the help of latest technologies such as machine learning, deep learning, simulation, complex decision-making, etc. On the other hand, IoT is considered as a network of connected physical devices that can process, run, and act on the given data. Apart from this, there also exists other key differences between them as illustrated in Table 1-1 [2] [3] [4].

The concept of DT was first officially coined by Michael Grieves in 2002 where he referred DT as a digital representation of a physical object, process or system [5]. The main role of a DT is to provide both dynamics and elements of how a given physical asset would operate in its entire life cycle. This is achieved by generating a digital counterpart of the corresponding physical asset. The generated digital counterpart is then utilized to carry out simulations related to the present and future state of the physical asset, study possible scenarios before they emerge, continuously gather, and analyse data on the physical asset to avoid unwanted circumstances, etc. These features when exploited results in the improvement towards a variety of application domains like retail [6], agriculture [7], manufacturing [8], healthcare [9], construction [10], etc. Furthermore, the massive rise in the utilization of DT is majorly because of the various benefits it provides such as improved productivity, reduced costs and downtime, optimized maintenance, and monitoring of physical assets, etc.

Table 1-1: Difference between DT and IoT.

Features	DT	IoT
Definition	Digital or virtual representation of a given physical object or process in real time.	Networked infrastructure of physical objects or things that are equipped with embedded systems like sensors, actuators, and other technologies to act on the given data.
Origin	Initially used by NASA in 2010 and first defined my Michael Grieves in 2002.	First used at Carnegie Mellon University in 1982 and was talked by Peter T. Lewis in 1985.
Types	Physical entity, virtual or digital entity, and the connection between them.	Cellular, low power wide area and mesh networks, RFID, Bluetooth, Wi-Fi.
Characteristics	Modularity, enhanced connectivity, reprogrammable, data homogenization.	Programmability, ubiquity, scalability, localization.
Application	Industries urban planning, manufacturing, healthcare, automotive.	Consumer applications, smart home devices, industries like healthcare, logistics, agriculture, manufacturing.

At present, various research studies have focused on the topic of DT. However, to the best of our knowledge, none of the research address the security issue that a DT system can face. Therefore, the goal of this paper is to propose an approach that studies and models the security aspect of DT in which the focus was done on a common application of DT as an in-patient monitoring system in the healthcare sector. Since a typical in-patient monitoring system compromise of several sub-components in different configurations, this paper therefore aims to propose a secure multi-layer DT architecture that captures the various attack and defence cases that can occur. This proposed model will be implemented using a probabilistic model checker called PRISM in which the specifications of the given system will be captured and verified by this software [11].

1.2. Problem Statement

While the rise in DT has brought forward advancements towards various industrial domains. Nevertheless, as in the case of any emerging technology, the extensive use of DT can also lead towards an increase in the execution of numerous security threats and attacks on to the actual physical entity which is derived from it. Attackers can capture and compromise the DT data throughout the DT layers and alter the system functionality. This data when obtained by the hacker can be used to serve as a blueprint to the physical system. This will in turn help the hacker in identifying the relevant components, behaviours, and interfaces that the physical object comprises of. The hacker can then study these components to gain an understanding of the internal view of the system and its vulnerable attack points. Furthermore, DT when acquired by the attacker can also serve as a potential platform for the execution of various backend system attacks. Therefore, it is important to consider the effectiveness as well as the security aspect of the DT system. However, there still exists an inevitable gap between the digital and actual system wherein lies the security vulnerabilities to the system. Therefore, the proposed work studies the various attacks and defences that are relevant to the DT layers and proposes a multi-layer security architecture. After which modelling of the proposed security architecture is performed using MDP and DTMC to analyse various potential attacks in each layer and their effect throughout the DT layers. To validate the proposed methodology, a case study in healthcare was chosen through which some of the security properties of the system were verified. Finally, results were presented which was then followed by conclusion and future work.

1.3. Thesis Objectives

The purpose of this thesis is to provide the following objectives:

- Provide a comprehensive review on DT components in each layer.
- Conduct an extensive study and analyse on the potential attacks that can compromise each DT layer.
- Propose a security architecture that covers all these layers.

- Develop modelling of the security architecture: security attacks and measures for DT.
- Implement a case study of security attacks and measures of DT in healthcare.

1.4. Thesis Organization

The rest of the thesis is organized as follows: Chapter 2 provides a brief background information about DT and its relevant literature review. Chapter 3 presents the modified conceptual DT model and talks in brief about it. Chapter 4 focuses on the security aspect of DT and presents the proposed multi-layer secure DT architecture. To model the security in the DT system, Chapter 5 talks in brief about the proposed method employed and its relevant background. Chapter 6 demonstrates the use case of DT in healthcare. Chapter 7 presents the experimental work. Chapter 8 presents the results and discussions obtained from the proposed work, and finally Chapter 9 concludes the thesis and provides possible future work.

Chapter 2. Background and Literature Review

This chapter provides a background on the concept of DT. It also presents a comprehensive literature review of the existing works related to DT and security modelling of IoT which is another related emerging technology.

2.1. Background

Despite the recent popularity received by the emergence of DT, it is important to note that the innovation of DT is not new. The application of DT can be traced back during the late 1960s at NASA's Apollo program [12]. In this program, NASA had to make two identical spacecrafts. The aircraft which is left on the earth is called the Twin. During the execution of this mission, the Twin was majorly used to simulate the space model so that it could help in precisely reflecting and predicting the status of the space vehicle. These simulated predictions would then help astronauts to make the correct decisions. Therefore, through this perspective, a Twin could be considered as a kind of a prototype model that helps to accurately reflect the status of a given product, system, or process (physical entity) in real-time via simulation.

However, the first idea of a DT was officially proposed during the year 2002 by Professor Michael Grieves at a Product Lifecycle Management (PLM) course that was held at the University of Michigan. He defined DT as a digital copy of one or more set of specific devices that can successfully illustrate a given physical entity. However, during that time, the concept of DT model was known as the Mirrored Space Model [13]. Later this concept was known as the Information Mirror Model [14]. Although, various terms were used to describe the DT model, however all these terms comprises of the same major elements that a typical DT model consists of such as the physical space, virtual space and the interface that exists between them. Then in 2011, Professor Michael Grieves, described the conceptual model of DT that comprises of the following three major components [15]:

1. The real space which consists of the actual physical product.
2. The virtual space which consists of the generated virtual model.
3. The data and information interface that exists between the real space and virtual space.

2.2. Types of Digital Twin

Based on its complexity and level of abstraction, DT can be classified into the following four major categories [16] [17]:

2.2.1 Component twin

A Component Twin refers to a Twin that represents a single key component that have a direct impact on the functionality and performance in the system. Apart from this, a Component Twin can also represent a component that is subject to high or jerky influences. Rotor, bulb, blade are examples of some of the possible physical components that can be derived to create a Component Twin.

2.2.2 Asset twin

An Asset Twin refers to a type of Twin that portrays the working of an entire asset which is nothing but a combination of individual components. To generate this Twin, any information received from a Component Twin, or a collection of Component Twins can be used. While the focus of a Component Twin is to improve the stability and robustness of the components, an Asset Twin on the other hand allows the user to study the operation of the entire system. The production of an Asset Twin helps to discover any possible improvement that a system can incorporate to enhance its performance. Motor, turbine, MRI machine are some of the examples of the possible physical components which can result into an Asset Twin.

2.2.3 System twin

Also commonly referred to as a Unit Twin. The System Twin is a type of a Twin that integrates individual Asset Twins. The purpose of this Twin is to help in analysing how well the system functions as a whole function when individual Asset Twins are combined. Aircraft, manufacturing conveyer belt, a set of equipment in an automatic car wash are some common examples of a System Twin.

2.2.4 Process twin

The purpose of a Process Twin is to provide a visual representation of an entire process. This process can include a combination of several systems that work together. It can also include a single object that carries out a certain process. For instance, a Process Twin in the manufacturing industry can be based on the data that is collected from every step of the process such as from the delivery of raw materials to their final transformation, moulding, designing, packaging, supplying, etc. Thus, this type of Twin can be commonly used in enhancing the business processes on a larger scale.

2.3. Characteristics

DT possesses certain characteristics that make them different from other technologies. Some of these characteristics are mentioned below [18] [19]:

2.3.1 Connectivity

One of the major characteristics of a DT is that it helps to enable connectivity between the physical entity and its corresponding digital counterpart. This connectivity is created by the sensors present on the physical entity that helps to collect, integrate, and communicate the collected data with the help of various technologies. Apart from this, the advent of DT has also helped in the establishment of connectivity between various products, organizations, and customers. For instance, the connectivity between the partners in an organization can be enhanced by allowing the members of the organization to check the status of a given physical entity by checking the generated DT of that entity and provide maintenance to its customers as and when required.

2.3.2 Reprogrammable nature

Another important characteristic of a DT is that it allows the physical entity to be reprogrammed in a certain manner via remote adjustments or through the help of artificial intelligence and predictive analysis. This reprogrammed physical entity can be then used as a basis for producing an improvised version of the initial physical entity. For instance, the DT of an engine can be reprogrammed to enhance its productivity and fuel efficiency.

2.3.3 Digital traces

Another characteristic of a DT is the fact that it can leave digital traces. A good example of the use of a digital trace is when the given machine crashes. In such a case, engineers can go back and check the traces of the DT to diagnose where the fault has occurred. Furthermore, these diagnoses can also be used by the manufacturer of these machines to enhance the design to avoid any type of malfunctions that can occur to the machine in the future.

2.3.4 Modularity

Modularity is another characteristic which can be achieved by the implementation of DT. The advantage of this characteristic is that it allows manufacturers to keep track of their machines and observe any potential areas for their improvements. Thus, when these machines are made modular using DT technology, it allows manufacturers to get an idea of which components perform can be replaced with better efficient components to enhance the overall manufacturing processes.

2.4. Benefits

Mentioned below are some of the major benefits that a DT provides:

2.4.1 Risk assessment

The emergence of a DT has enabled organizations to test and validate their respective products before producing it into the real world. Through the creation of a virtual replica of the physical product, a DT lets engineers to recognize any possible failures that could occur on the actual product itself. Apart from this, engineers can also produce unexpected scenarios by disrupting any operation of the system. This helps them to further examine the reaction of the system and study any corresponding mitigation strategies proposed. Thus, DT can help immensely in the improvement of risk assessment and in the enhancement of the products reliability.

2.4.2 Predictive maintenance

One of the important benefits of a DT is that it helps to solve problems in advance. That is, it facilitates the ability to perform predictive maintenance for a given physical entity. This is possible because of the sensors present in the DT system which help to gather large amount of data in real-time. The analysis of this collected data can be proactively used to examine any faults present within the system. Any fault present in the system will be then received as a report to the human operations who would then address the problem faced in a timely manner. When a component of the physical entity can be replaced before being broken, then it would help the manufactures to avoid any serious damage, unnecessary downtime, and expenditure. Thus, the utilization of DT greatly helps in enhancing the production line efficiency and in the reduction of the maintenance costs.

2.4.3 Production time

One of the major goals for any organization is to be on the market faster than their competitors. However, this goal is not easy to achieve as it all depends on the tedious steps and constant changes that occur during the production processes of a product. Nonetheless, the adaptation of DT helps to carry out the life cycle of a particular product in the virtual environment, where all types of improvements can be executed in a faster and efficient manner. Moreover, the virtual prototype created in the DT system helps to validate how the physical product would function. Thus, the utilization of DT helps organizations to significantly reduces the production time, optimizes the efficiency and development time for a given product.

2.4.4 Implement decisions in real-time

The implementation of a DT enables decision makers to quickly understand the implications of any changes done to a physical entity at any given point in time. For instance, if a material for a given physical entity is changed, what kind of impact would it have on the design model, project plan model, overall cost, etc. Thus, the emergence of a DT can greatly enable organizations to execute simulations to answer a particular "what if" queries. Based on these queries, any kind of adjustments can be done rather than going through the entire process of creating a physical model itself.

2.5. Applications

The emergence of DT is transforming the way tasks are performed across different industries. Knowing these applications can help businesses implement DT into their processes. Therefore, mentioned below are some of the major applications of DT in the following industries:

2.5.1 Manufacturing

DT can be majorly used in the manufacturing sector. This is because they play a significant role in the way products can be designed, manufactured, and maintained. That is, numerous sensors are placed throughout the manufacturing process. The role of these sensors is to collect the various types of environmental and operational data of the machine as well as the work that is being executed. DT then uses this data to generate a virtual replica of the near real-time occurrences. In addition to this, they can also help in enabling the manufacturing companies to achieve a digital footprint of all their products throughout their entire product life cycle that ranges from the time they are designed to the final process of their development [20].

2.5.2 Healthcare

Healthcare is another sector that can majorly benefit from the implementation of DT. The idea of utilizing DT in the healthcare has been actively used in equipment or product prognostics [21]. Furthermore, the utilization of DT in the healthcare sector was made possible due to the vast availability of technologies that act as a personalized product for the patients. The role of these technologies is to continuously record the different vital signs of the patients throughout their lifecycle. This process can be used in a form of a personal healthcare management system which can be useful especially for the elderly patients [22]. These data can be then used to ultimately generate a detailed Virtual Twin of the given patient itself. Apart from this, DT can also be used to compare the individual records of the patients with the entire population to find health related patterns which would help to improve patient diagnosis.

2.5.3 Automobile

The concept of DT can also be used in the automobile sector. They can be implemented in the automobile industry through the help of various existing data that helps to facilitate their process and lower the marginal costs. That is, a DT in the automobile industry can be used to capture both the behavioural and operational data of a given vehicle. This data then can be used to analyse and improve the overall performance of the vehicle as well as the features connected to it [23].

2.5.4 Retail

Meeting customer expectations and improving their experience are some of the key goals that needs to be achieved for any retail sector. The implementation of a DT can help achieve these goals [24]. For instance, retailers can help deliver the ideal fashion clothing products to their customers based on their respective DT models. This would thereby help to tremendously improve the overall businesses in the retail sector. Apart from this DT can also be used to improve instore planning in a more efficient manner.

2.5.5 Energy

DT can be employed in the energy sector to enhance the optimization and maintenance process of the physical systems, assets, and process present. The implementation of DT for power plants is one example of how DT can be used in the energy sector [25].

2.6. Related Work

In this subsection, various research papers that focuses on the following topics were analysed:

2.6.1 Digital Twin

Currently, several research papers exist that provides a comprehensive survey on DT which mainly focuses on its application across a variety of domains. Barricelli et al. [26] in their paper presented a comprehensive survey that comprises of various definitions of DT and other related concepts. Next, they analysed the selected papers to identify the major characteristics of a typical DT and presented the various application domains of DT. Apart from this, their paper also provided a brief overview of design implications that were derived from their study performed. Finally, major open issues and challenges that requires to be addressed and studied for the successful implementation of DT were presented.

Similar to the work done above, Fuller et al. [27] also presented a comprehensive review of DT where they first defined what a DT is and tackled some of the common misconceptions that are associated with the current and previous definitions of DT. Next, they discussed the challenges faced by DT, and investigate on the key emerging technologies of DT while focusing on each of their histories as well. After which they presented a categorical review of recent papers in which they categorized different papers that reflect different areas and their current state of research. Manufacturing, healthcare, smart cities, etc., are some of the research areas that their paper mentioned. Lastly, they concluded their paper by providing an evaluation of the enabling technologies that can be employed, the challenges and possible future directions that can be faced by implementing DT.

Apart from comprehensive surveys, an application framework for the entire product management lifecycle was also proposed by the work done by Zheng et al. [28]. The proposed application framework used the key concept and major characteristics of DT that were extracted from both broad and narrow sense. The framework comprised of three major parts: the physical space, the information processing layer and finally the virtual space. In the physical space, they discussed in detail regarding the total element information perception technology of production. Whereas the information processing layer was further divided into three major function modules that included the data storage, data processing, and data mapping. Ultimately, for the virtual space their paper discussed about the implementation process of the full parametric virtual modelling and the construction idea for DT application subsystems. Lastly, their paper concluded by studying a DT use case of a welding production line.

On the other hand, to highlight the impact and importance of DT in various industries, Uhlenkamp et al. [29] presents a systematic classification related to the various DT applications. First, they provided a brief literature review related to DT which helped to demonstrate an overview on the development of DT that has occurred throughout the last few years. Next, they focused on the diverse applications of DT. Their paper focuses on the three major application use cases of DT. Later, they also analysed their conceptual background, the targeted problem and accordingly implemented a suitable use case. Based on the results of their analysis obtained, their paper claimed that they could categorize the mention DT applications based on seven dimensions that include distinctions of goals, focused users, life cycle phases, system levels, data sources, authenticity, and data exchange levels.

Apart from the survey papers on DT, there also exists various papers that focuses on one key application of DT. Lu et al. [30] in their work reviewed the recent development of DT technologies in manufacturing systems and processes in which they first provided a brief overview on the concept of DT which is then followed by an in-depth discussion regarding the implication of a DT derived smart manufacturing system. Next, they highlighted on how DT will help to tremendously enhance the future of the manufacturing sector. Later, they also provided a detailed DT reference model and key enabling technologies that help in the development of a DT driven smart manufacturing solution. Lastly, possible research challenges and future research directions were provided in their paper. The application of DT in the supply chain sector was also talked about by the work in. [31]. Here the authors presented some of the key research issues of DT which needs to be tackled. In addition to this, appropriate future research directions have been also suggested by the authors as well.

Smart automotive is another domain that has benefited from DT implementation. Thus, Bhatti et al. [32] presented a review regarding the application of DT in smart electric vehicles. First, they provided a conceptual background of DT in their paper. Later they shifted their focus to give attention to the contributions of DT in smart vehicle systems. Based on this, they have classified their review into various domains that exists within the smart vehicle systems. Autonomous navigation control, advanced driver assistance systems, vehicle health monitoring, battery management systems, vehicle power electronics, and electrical power drive systems are some examples of the domains that they covered. An in-depth discussion of each of these domains were covered in their paper. Lastly, they concluded their work by talking about the challenges and future scope of implementing DT in the automotive sector.

To demonstrate the benefit of DT in the healthcare sector, the work done by Rivera et al. [33] presented their work that talks about the application of DT in precision medicine. Their main contributions can be divided into two major categories. First, they described their initial ideas for a reference model that would enhance the capabilities that DT provides to design a smart and flexible software system for the healthcare sector. The designed system is expected to reduce complexity and assist in the planning and decision-making processes during the process of implementing medical treatments to the patients by the healthcare professionals. In addition to this, they also elaborated on the internal structure for DT that would help to support precision medicine techniques.

In addition to the papers mentioned above, several other published works proposed a real-life application of DT. For instance, in. [34], the authors demonstrated how DT can help improve in the advancement towards smart farming. They also talked about the concept of DT and provided a relevant typology for it. Later they built a conceptual framework for designing and implementing DT in smart farming. Their framework was built through an analysis of the literature review performed regarding the concept of DT. Apart from this, they presented a review on how DT can be employed in the smart farming sector. Lastly, to validate their proposed conceptual framework for smart farming, five major case studies from the Europe IoF2020 project were employed. Arable farming, dairy farming, greenhouse horticulture, organic vegetable farming, and livestock farming were the chosen five case studies for their work.

In terms of utilizing DT in the healthcare sector, Laaki et al. [35] demonstrated a prototype of autonomous surgery that harnesses IoT and industry 4.0 connectivity to generate a DT of a patient. They proposed a remote surgery application via a mobile network. The proposed prototype utilizes a robotic arm, Virtual Reality with a 4G environment, to execute a precision surgery. In addition to this, they also discussed the possible challenges faced when integrating the prototype with a DT. Lastly, they investigated on some of the advancements in Artificial Intelligence and industry 4.0 and suggested on how one can ease the challenges of connectivity, integration, and multidisciplinary research.

Another use case of DT in smart grid can be seen from the work presented by Sivalingam et al. [36] that reviewed and presented a case study that examines the wind farm use and energy consumption in the case of smart grid. Their work presents some of the challenges present with the reliability of power consumption and the general maintenance of wind turbines. Then they proposed a novel methodology that uses IoT sensors in combination of data analytics within a DT environment to precisely perform predictive maintenance of the given wind turbines. Finally, to further enhance their proposed work, the authors presented possible future work which can be implemented.

To provide a safer driving experience for the road users, Chen et al. [37] in their work demonstrates on how a DT can be used for cars and traffic management. Their research first explored the various challenges that can be faced while driving. Then they demonstrated a framework that utilizes DT in addition to various learning algorithms to monitor and analyse feedback based on user behaviour. The purpose of these algorithms is to facilitate a real-time digital behavioural Twin of a driver and provide possible warnings and instructions on how-to drive-in order to reduce the occurrences of possible risks faced.

In terms of the security aspect of DT, the work done by Al-Ali et al. [38] mentioned about the presence of a security layer in their proposed six-layer end-to-end conceptual model in which they explained in detail the functionality of each layer. That is, their proposed conceptual model comprises of six major layers: the physical space layer, the communication network layer, virtual space layer, and the application layer. In addition to this, their paper also talked about the need of a security layer that needs which would overlap throughout the entire architectural layers. However, they did not focus in detail on the security aspects of DT. Thus, based on the extensive research conducted, a lack of research that focuses on the security aspect of DT were found. Furthermore, since DT architecture is similar to the architecture of IoT. Therefore, the upcoming subsection presents relevant research papers that focuses on the different ways of modelling the security aspect of IoT.

2.6.2 IoT security

Contrary to the previous case for DT, there exists several papers that focuses on the security aspects present in IoT. However, in this section, the focus of the thesis was narrowed down to only those papers that addresses the security aspect of IoT. One example of such a work is the paper presented by Potrino et al. [39]. The authors in this paper analysed and modelled a novel IoT security system. The context of their work is based on a typical IoT system in the presence of lightweight sensor and actuator nodes that exchange messages using the Message Queue Telemetry Transport (MQTT) protocol. The goal of their work is to help to mitigate the security challenges faced by this protocol, especially the possible DoS attacks that can occur in the application. Their system is based upon the utilization of the host Intrusion Detection System (IDS) that applies a threshold depending upon the packet discarding policy through the various topics established via the MQTT.

On the other hand, the work presented by Ge et al. [40] aims to enhance the IoT security by providing a framework for modelling, assessing and also by providing a formal definition of the framework. Their proposed framework comprises of five levels which are data processing, security model generation, security visualization, security analysis, and model updates. With their framework, any possible attack scenarios in the IoT system can be captured. Furthermore, their proposed framework also helped to analyse IoT security via the defined security metrices and assess the strength of different defence mechanisms. In terms of evaluation, their framework was evaluated using three different case studies: smart home, wearable healthcare monitoring and environment monitoring. The obtained analysis results were used to demonstrate the capabilities of the proposed framework in capturing the possible attack paths and mitigating the effect of the attacks.

In [41], the authors introduced a lightweight biometric based remote user authentication and key agreement scheme that they proposed in order to provide a secure access for the IoT services. Their proposed uses a lightweight hash operation as well as XOR operation. Through the security analysis performed, it was shown that their work could withstand various security attacks. As part of the experimentation, they utilized a formal verification tool called AVISPA to prove the security of the protocol in the vicinity of a potential attacker.

Zahra et al. [42] in their work evaluated the effect of the Shibboleth protocol in a Cloud-IoT network in order to ensure a secure data outsourcing and access. In their work, they added the Shibboleth control protocol between the fog node and fog client to improve and provide secure communication between them. In terms of experimentation, they have verified the protocol using the high-level Petri net tool. Apart from this, the Z3 SMT solver was also to study the rules of the information flow that helped in proving the correctness of the given system against the provided security properties.

At present, there exists several variable security risks associated with the IoT system. Therefore, in order to formally and quantitatively analyse these risks, Mohsin et al. [43] presented an IoTRiskAnalyzer framework. Their work utilizes a using probabilistic model checking approach in which their framework collects vulnerability scores, candidate IoT configurations, as well as the capabilities of the attacker as input. After which, a system and threat model is generated to compute the attack possibility and its associated cost for every configuration. Through the evaluation of their work, the authors found out that the proposed framework is efficient and automatically prioritizes the input configurations based upon risk exposure.

In terms of research that mainly focuses on security analysis in IoT applications, several security analyses of IoT in the healthcare domain exists. For instance, in [44] the authors presented an efficient and strong authentication protocol for the secure access of patient data. Their work was for the healthcare applications that were based on the Cloud-IoT network. Through the experimentation performed using the AVISPA tool, their proposed work proved to be secure against possible attack (both passive and active).

To address the security and privacy challenge related to medical data for the IoM, Deebak et al. [45] in their work provided Secure and Anonymous Biometric Based User Authentication Scheme (SAB-UAS) to ensure secure communication in healthcare applications. Upon the performance analysis conducted it was shown that the proposed work also proved to have high-security features. In terms of experimentation, their work used the NS3 simulator for analysing the network parameters. The obtained results demonstrated that their work showed superior results in terms of the end-to-end delay, packet delivery, routing overhead, and throughput rates.

Merabet et al. [46] in their work proposed three novel lightweight authentication protocols for the healthcare applications that are based upon the IoT systems. Amongst the three protocols, two of them were shown to be suitable for M2C communication, whereas the last one was suitable for M2M communication. As part of the experimentation, to formally verify their work, the AVISPA and ProVerif automated tools were utilized. By using these tools, it was shown that the proposed work satisfied all the security requirements.

In order to provide secure and private communication between patient and their devices, Theera-Umpon et al. [47] in their work proposed a security protocol that manages the vulnerabilities that exists in the communication between implantable medical devices and other devices. Their work also encrypted the data and the communication process to eradicate the possibility of confidential data being leaked. Verification of their work showed that their work provided the safety and security in wireless communication.

As illustrated in Table 2-1, current papers only focus on the security aspect of IoT. Therefore, this thesis presents a novel method that focuses on the modelling the security aspect of DT.

Table 2-1: Summary on the Various Methods Based on IoT Security Modelling.

Papers	Aim	Method Employed	Drawback
Potrino et al. [39]	Evaluated and modeled IoT security system.	Event driven simulator called Omnet++.	Focused on enhancing the security at the application level.
Ge et al. [40]	Proposed a framework to model and analyze IoT security.	Graphical security model and a security evaluator.	Did not provide any simulations and experiments for their work.
Dhillon and Kalra [41]	Proposed a lightweight biometric based user authentication and key agreement scheme.	Formal verification tool called AVISPA.	Focused on enhancing the security at the application level.
Zahra et al. [42]	To study the security of data communication present in Fog-IoT systems.	High Level Petri nets was used and Z3 SMT solver	Focused on outsourcing and providing secure data access.
Mohsin et al. [43]	Proposed framework to analyze and risks	Probabilistic model checking tool called PRISM.	Proposed framework does not take defence mechanisms into account.

Dhillon and Kalra. [44]	Presented a strong and efficient secure authentication protocol	Formal verification tool called AVISPA.	Focused on security issues that arises due to storing of data.
Deebak et al. [45]	Proposed a Secure and Anonymous Biometric User Authentication Scheme	NS3 simulator	Focused on providing a secure communication in the system.
Merabet et al. [46]	Proposed three lightweight authentication protocol	AVISPA and ProVerif automated tools	Majorly focused on providing a secure M2C and M2M communication.
Theera-Umpon et al. [47]	Proposed a security protocol	Communication Sequential Process (CSP)	Mainly focused on providing a secure communication between IMDs and other devices.

Chapter 3. Digital Twin Architecture

This chapter presents various definition of DT. Apart from this, it also presents the modified conceptual DT model and explains it in brief.

3.1. Digital Twin Definition

At present several research papers have defined the concept of DT in a variety of ways. Table 3-1 depicts some of these definitions.

Table 3-1: Various Definitions of Digital Twin.

Year	Authors	Definition
2020	Rasheed et al. [48]	"*Digital Twin can be defined as a virtual representation of a physical asset enabled through data and simulators for real-time prediction, optimization, monitoring, controlling, and improved decision making".*
2019	Madni et al. [49]	*"A Digital Twin is a virtual instance of a physical system (Twin) that is continually updated with the latter's performance, maintenance, and health status data throughout the physical system's life cycle"*
2019	Zheng et al. [28]	*"A Digital Twin is a set of virtual information that fully describes a potential or actual physical production from the micro atomic level to the macro geometrical level."*

3.2. Digital Twin Architecture

However, in terms of understanding the working of the DT, an architecture of DT is required. Therefore, a generic DT model is derived from the work conducted by Al-Ali et al [38]. The purpose of this model is to clearly depict how a DT would operate irrespective of the industrial domain it is used in. As depicted in Figure 3-1, the proposed model includes five major layers: the physical domain layer, communication network layer, virtual domain layer, application layer and the security layer. These layers are grouped in accordance with the common features it possesses. The virtual domain layer can be further classified into the data acquisition unit, data aggregation and modeling, and the data analysis and visualization sub-layers.

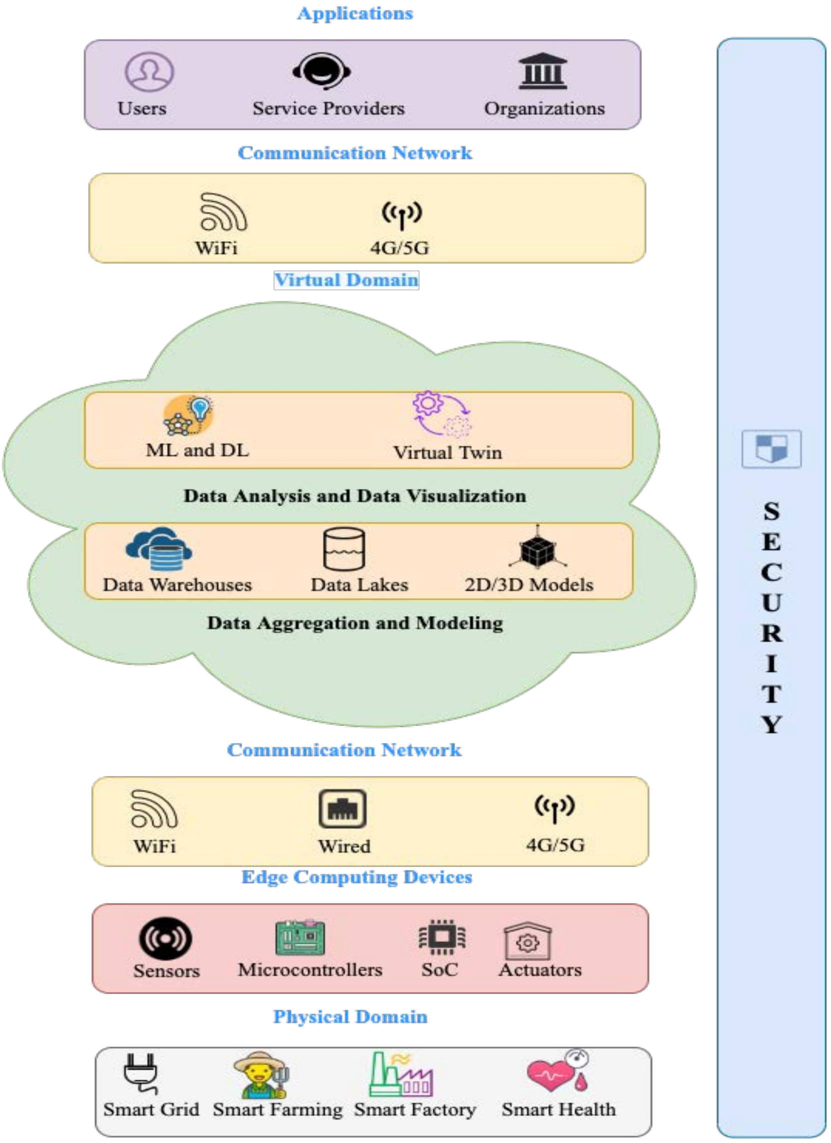


Figure 3-1: Modified Conceptual Digital Twin Model [38].

3.2.1 Physical domain layer

In the context of a DT, any given product, process, or system comprises of two major domains: the physical domain and the virtual domain. As illustrated in Figure 3-1, the physical world comprises of various smart applications like smart healthcare, smart agriculture, smart building etc. However, the smooth functioning of these smart applications requires the involvement of one or more sensors that come equipped with a given physical entity itself. The role of these sensors is to help gather operational data and data related to its surrounding environment. This data is then converted from non-electrical signal to electrical signal by the sensors itself. The collected data can be anything ranging from a simple temperature measurement to a complex video feed. It could also exist in the form of location, humidity, sound, different vital measurements of the human body, etc. Currently, there exists a variety of IoT sensors that can be used according to a given application. The main advantage of these sensors is that they are cheaper, smaller in size and consume less power. Mentioned below are some of major types of sensors that are often used for the implementation of various smart applications:

3.2.1.1 Human vital sign sensors

IoT sensors can greatly benefit the healthcare sector. That is, these sensors can be used to measure and monitor different medical parameters for a given patient even if they are not in the hospital or if they are all alone by themselves. Apart from this, these sensors can also be used to provide feedback in real-time to the doctor, nurse, caretaker, respective families of the patient, or even the patient himself. At present, various wearables devices are available in the market that are equipped with medical sensors who are capable measuring different vital information such as blood pressure, blood sugar, heart rate, etc. [50]. Smart watches, monitoring patches, wristbands and the like are some of the major examples of the major wearable devices. Famous companies like Apple [51], Samsung [52], etc., have designed their own smart watches and fitness trackers that provides various useful features like heart rate monitor, connectivity with smartphone, blood pressure monitor, physical activity tracker for activities like cycling, swimming, running, etc.

3.2.1.2 Radio Frequency Identification (RFID)

RFID is a type of sensor that uses radio waves to read and capture information that is stored on a tag which is attached to a given object. This tag can be read from several feet away and it does not require to be within the direct line-of-sight of the reader to be tracked. This is the major advantage that it has as compared to the barcode asset tracking system. In general, a typical RFID system comprises of three major components: the tag/label, reader, and an antenna. Furthermore, the RFID tag comprises of the following major types:

1. Active Tag: This type of RFID tag comprises of its own power source. In terms of range, it has a broadcast range of up to 100 meters.
2. Passive Tag: This type of RFID tag does not comprise of its own power source [53]. However, it is powered by a reader. In terms of range, it has a read range from near contact to up to 25 meters.
3. Semi-Passive Tag: This type of RFID tag comprises of a battery. However, it communicates with a reader using a backscatter just like a passive tag without a battery that helps to provide a longer read range as compared to the traditional passive RFID tags.

Regardless of the RFID tag chosen, the RFID in general comprises of an integrated circuit and an antenna which can be used to transmit the data to the RFID reader or the interrogator. The role of the reader is to convert the radio waves into a useful form of data. The information acquired from the tags is then transmitted via a communication interface to a host computer system. Here the data can be stored onto a database so that it can be analysed later whenever required. Figure 3-2 depicts an example of a typical RFID system.

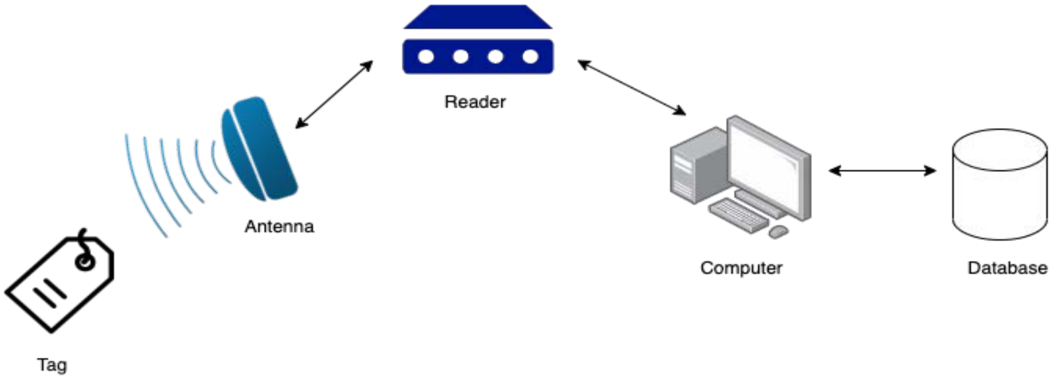


Figure 3-2: Typical RFID System.

3.2.1.3 Environmental sensors

The purpose of environmental sensors is to sense parameters present in the physical environment. These parameters could be humidity, pressure, temperature, air pollution, water pollution, etc. The temperature and the pressure parameters can be measured using a thermometer and barometer. Whereas air quality can be measured by the help of sensors that help to sense the presence of gases or any other matter present in the air.

3.2.1.4 Chemical sensors

Chemical sensors are measurement devices whose role is to detect any traces of biochemical and chemical substances [54]. Once these substances are detected, these sensors then convert a physical or chemical property of the detected chemical substances into measurable signals. They can be employed to countless domains like automotive, medical, nanotechnology, in home detection systems as carbon monoxide detectors, etc. Various chemical sensors exist that are specifically designed to perform a particular task. Regardless of the tasks they perform, most of the chemical sensor comprises of receptors and transducers. The role of the receptor is to transform the chemical information into a form of energy. Whereas the role of the transducer is to transform this energy into a measurable signal.

3.2.1.5 Actuators

Apart from sensors, the physical domain layer also comprises of actuators that operates in the opposite way as compared to how sensors behave. That is, an actuator takes an electrical input and convert it into some form of a physical action. In general, an actuator can be classified according to the type of operation it can perform. Mentioned below are the four major types of an actuator:

1. The hydraulic actuator is a type of an actuator that operates through the utilization of a fluid-filled cylinder with a piston that is suspended at the centre. These actuators often generate linear movements, and a spring is fastened to one end as a part of the return motion. They can be widely used in various exercise equipment such as car transport carries, steppers, etc.
2. The pneumatic actuator is a type of an actuator that is commonly preferred to when it comes for machine motion. They utilize pressurized gases to generate a mechanical movement. Several companies utilize these actuators because they can generate a highly precise motion especially when the machine is in the state of being started or stopped. Pressure sensors, pneumatic mailing systems, grippers, tie-rod cylinders, etc., are some examples of equipment that utilizes pneumatic actuators.

3. The electric actuators as the name suggests relies on electricity for their operation. It is like the pneumatic actuators in the sense that it can also create precise motion. This is because the electrical power is constant in nature. Electric cars, manufacturing machines, robotics equipment, etc., are some of the well-known examples of equipment that utilizes electric actuators. In general, the electric actuators can be majorly classified into electrohydraulic and electromechanical actuators. Where the electrohydraulic actuator is a type of actuator that is also powered electrically. However, it gives movement to a hydraulic accumulator. This accumulator then provides the required force for movement. On the other hand, an electromechanical actuator is a type of actuator that helps to transform the electrical signals into linear or rotary movements or a combination of both.
4. Magnetic and thermal actuators often compromise of shape memory allows that can be heated to produce movement. The motion of magnetic or thermal actuators majorly occurs from the Joule effect. However, it can also arise when a coil is kept in a static magnetic field. The purpose of this magnetic field is to generate a constant motion that is most referred to as the Laplace-Lorentz force. One of the major advantages of these kind of actuators is that it can produce a wide variety of powerful motion while still being lightweight.

Once the data generated from the sensors is acquired, it is later sent to the microcontrollers. The role of the microcontroller is to further process the collected data from the sensors and then send it back to the actuators to execute the necessary actions. Furthermore, selected operational and environmental parameters would be transmitted to the communication network layer to further process as well as store the data in the above layers. These microcontrollers are equipped with large store, high speed CPU, large number of analog and digital ports, etc. In addition to this, they also support the major wired communication ports such as SPI, RS232, I2C, CAN, and USB. They also support various wireless access points such as Bluetooth, GPRS, Ethernet ports, etc [38]. Furthermore, they also often equipped with numerous proprietary ports to support various external devices such as camera, keypads, LCD, GPS, etc.

Currently, there exists a variety of microcontrollers that are present in the market. Some of the major microcontrollers are mentioned below:

3.2.1.6 Raspberry pi 4 model B

It is the latest version of the low-cost computer. In its cheapest form as it does not have a case and is just simply a tiny electronic board that is just the size of a typical credit card. In terms of cost, it starts from a minimum cost of $35. However, it also is available at a cost of $45 and $55 respectively, depending upon the specifications required. Over the years, it has gained a lot of popularity. This is because of its affordable price, user-friendly design, and compact size. The $55 Raspberry Pi 4 Model B is considered as the most powerful Raspberry Pi present yet as it provides a tremendous increase in the processing power, video output, peripheral connectivity. In addition to this, it also provides numerous other features like power over USB Type-C, Ethernet connectivity, video output that can handle 4K monitors at a time, USB 3.0 ports, etc. Figure 3-3 illustrates how a Raspberry Pi 4 Model B looks like.

Figure 3-3: Raspberry Pi 4 Model B.

3.2.1.7 Particle photon

The Particle Photon is a tiny $19 Wi-Fi development kit that can be used in creating connected products as well as projects for IoT. It is easy to use, powerful and is connected to the cloud. It consists of a powerful STM32 ARM Cortex M3 microcontroller and a Broadcom BCM43362 Wi-Fi chip as its connection to the internet. Apart from this, it also comprises of 18 mixed GPIO pins and a web-based IDE**.** The RGB LED present in the Photon and the two buttons (one for setup and other for reset) can be used to switch between the different modes to help the user to debug their respective projects. Figure 3-4 illustrates how a Particle Photon looks like.

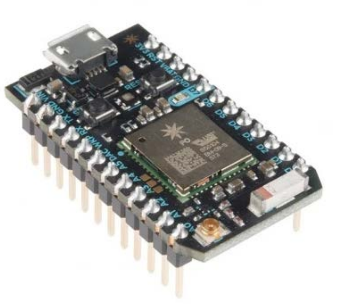


Figure 3-4: Particle Photon.

3.2.1.8 ESP32

It is a series of low power and low-cost System on a Chip (SoC) microcontrollers that provides connectivity features like dual-mode Bluetooth and integrated Wi-Fi. It was developed and created by a Shanghai-based Chinese company called Espressif Systems as a successor to the ESP8266 microcontroller that is slightly expensive but much more powerful SoC. It employs a Xtensa LX6 microprocessor in both single-core and dual-core variations that comprises of power amplifier, RF balun, filters, power-management modules, built-in antenna switches and low-noise receive amplifier. Mostly all the chips in the ESP32 series are dual core except for the ESP320S0WD which is single core. Although the ESP32 has around a total of 48 GPIO pins. However, only 25 of them can be used as pin headers on both of sides of the development boards. These pins can be used for numerous peripheral purposes such as for SPI, I2C, I2S interfaces, UART interfaces, ADC channels, DAC channels, touch pads, etc. It also comprises of two buttons: one is reset button to reset it and the other is the boot button which can be used to download new programs. Figure 3-5 illustrates how an ESP32 looks like.

Figure 3-5: ESP32.

3.2.1.9 Microchip DM990004

The Microchip DM990004 also known as the IoT Ethernet Kit is powered by AWS IoT. It uses an Ethernet LAN8740A that comprises of various features like deterministic loop back delay, ensuring real-time system performance and also cable diagnostics that help to reduce the network installation costs. This kit is controlled by a PIC32MZ EF 32-bit microcontroller which provides a 2 MB flash and enough space to store applications. In addition to this, it also provides a user experience with a pre-installed firmware which allows communication with the AWS IoT. Connection to the cloud is achieved by the utilizing the AWS IoT service provided. The AWS IoT is basically a cloud platform that allows connected devices to easily interact with the cloud applications and other devices in a secure manner. Apart from this, the utilization of AWS IoT service can also help applications to keep track of and communicate with the other devices even if they are not connected. In terms of sensors, several sensors can be plugged into this microcontroller through the help of the MikroElektronka mikroBUS™ which allows the prototyping of numerous IoT Proof of Concepts (PoC). Figure 3-6 illustrates how a Microchip DM990004 looks like.

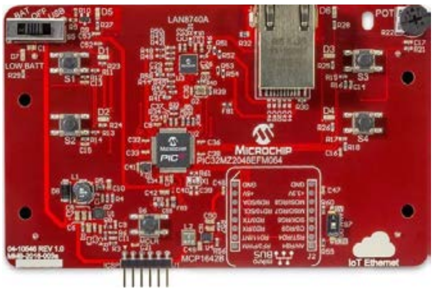


Figure 3-6: Microchip DM990004.

3.2.1.10 Intel quark D2000

The D2000 is a 3.3 V board that has an operating range between 2.0-3.3 V. With on-board regulations, it can be powered via a USB connector. Alternatively, there are also screw terminals present for external supply. In addition to this, all of the I/O is 3.3 V, and it comprises of a good deal of I/O functionality. However, it comes with the cost of multiplexing. That is, 25 I/O pins can be configured as GPIO or for other functions like SPI, UART, JTAG, I2C, etc. In this regard, it comprises of four user mode configurations. Besides the user modes, there is also a pin test mode that it comprises of. The GPIO have programmable drive strength of 12 mA and 16 mA modes and integrated pull-ups. Also, in terms of analog input, there are up to 19 analog inputs as ADCs or comparators. These ADC inputs are programmable in nature and has a 6,8,10,12-bit resolution. The analog comparators are either fast speed or slow speed, low-power with wake-capabilities. Figure 3-7 illustrates how an Intel Quark D2000 looks like.

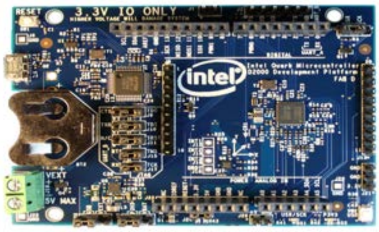


Figure 3-7: Intel Quark D2000.

Although there exist numerous features in the various microcontroller that were discussed above. However, to narrow down the topic of this section, a summary of the common features in a tabular form is presented. Therefore, Table 3-2 illustrates a visual comparison between the various microcontrollers in terms of CPU core, speed, RAM, SD card, networking, Bluetooth, GPIO, ADC, type of communication, cloud connectivity, and the price of each of the microcontrollers. As seen from the table, the most powerful microcontroller is the Raspberry Pi 4 Model B that provides a RAM up to 4 GB. It is also the fastest as compared to the others, as it can operate with a speed of 1.5 GHz. However, in terms of the price, the ESP32 is the cheapest with a unit price of $4 whereas the most expensive is the microchip DM990004 at $99.

Table 3-2: Comparison of Various Microcontrollers.

Features	Raspberry Pi 4 Model B [55]	Particle Photon [56]	ESP32 [57]	Microchip DM990004 [58]	Intel Quark D2000 [59]
CPU Core	Broadcom 2711 64-bits	STM32F205 with ARM Cortex M3	Tensilica Xtensa Dual-Core 32-bits	PIC32MZ2064DAH176	Quark D2000 with x86

					Pentium processor
Speed	1.5 GHz	120 MHz	240 MHz	200 MHz	32 MHz
RAM	1 GB, 2 GB, 4 GB	128 KB	512 KB SRAM	32640 KB	8 KB SRAM
SD Card	Up to 32 GB	Add-On	Add-On	Add-On	Add-On
Networking	Ethernet, Wireless	Ethernet Add-On, Wireless	Ethernet, Wireless	Ethernet	Add-On
Bluetooth	Built-in	Add-On	Built-in	Add-On	Add-On
GPIO	40	Up to 15	36	Up to 120	25
ADC	Add-On	7	12	Up to 45	19
Communication	I2C, SPI, CAN Add-On	RS232, I2C, SPI, CAN	I2C, SPI, I2S, CAN	RS232, I2C, SPI, CAN	RS232, I2C, SPI
Cloud Connectivity	Can be turned into a personal cloud	Google Cloud based	Google Cloud based	Amazon cloud based	Wind River cloud based
Price	$35, $45, $55	$19	$4	$99	$15

3.2.2 Communication network layer

The communication network layer acts as a bridge that connects between the physical domain layer and the virtual domain layer. The role of this layer is to effectively transmit or receive the data collected by the DT as well as the results collected by the processing of the DT. The collected information is then passed on to the higher layers where further processing and analysis can be performed. Since the virtual domain of the proposed architecture may not necessarily be present in the same geographical location as the physical domain, the implementation of large area wireless networks is therefore necessary. Thus, various mobile communication technologies and satellite communication networks can be implemented to ensure seamless communication between the two domains. In the context of mobile communication technology, various mobile communication networks are present. Therefore, mentioned below are the different types of mobile communication networks:

3.2.2.1 1G, 2G, and 3G cellular network

The first-generation mobile networks first emerged in Japan in the year 1979. After which it emerged in the United States in 1980 and the UK in 1985. This cellular network was based on an analogue technology which uses Frequency Division Multiple Access (FDMA) modulation. The utilization of this network successfully provides a channel capacity of 30KHz and a speed of 2.4kbps. However, the drawback of this network was that it was only restricted to voice calls, suffered from reliability and signal inference issues, and had limited security protections against hackers.

The first major upgrade received by cell phones was when they switched from 1G cellular network to 2G cellular network. One of the main differences between these two cellular networks was that the radio signals that were used by the 1G cellular network were analog in nature. Whereas in the case of 2G cellular network these radio signals were digital in nature. Furthermore, as compared to 1G cellular network, the 2G cellular network not only provide voice calls, but it also provided other services like Multimedia Messaging Services (MMS) and Short Message Services (SMS). Lastly, the 2G cellular network also offered other features like high data rates, bandwidth, and a secure and reliable communication channel.

Later, the 3G cellular network was first introduced commercially in 2001. It was the first generation of cellular network technology that provided numerous services which are used today such as web browsing, email, video downloading, etc. The major goal of this cellular network was to support various types of applications, increase data transmission at a lower cost, facilitate greater voice and data capacity, etc. In terms of data rates, 3G networks provides a data rate of 2Mbps for stationary users, 384 kbps for low-mobility users, and finally 144 kbps for high-mobility users. Apart from this a bandwidth of 20 MHz is provided by this network.

However, to ensure the smooth operation of DT, certain network requirements needed to be fulfilled. In addition to this, high bandwidth and high data rate were also required to transmit the vast amount of real-time collected into the virtual domain layer. Furthermore, a very low latency is also required to ensure the smooth transfer of the data from the virtual domain layer back to the physical domain layer. That is, it is important to ensure a minimal delay when the data is transferred back to the physical domain layer after being examined and analysed by various software that are present in the virtual domain layer. Another major factor to consider during the implementation of DT is the scalable and network capacity of the selected communication technology. That is, in the context of network capacity, this means that a single base station should be capable enough to hold simultaneously as many connections as possible. Whereas, in the context of scalable capacity, this means that an increase in the hardware components of the communication systems should lead to a significant improvement in the network so that the DT can collect as much data as possible and spread it across the entire product, system or process depending on the given domain.

The above-mentioned mobile communication networks do not sufficiently meet the requirements for implementing a DT. Thus, 4G, 5G and even the upcoming 6G networks can be considered for its implementation. Table 3-3 illustrates a summary of the differences between the different communication networks.

3.2.2.2 4G cellular network

This generation of cellular network was only made possible practically because of the advancements in mobile technology that occurred during that time. It was the first generation of large area wireless network that allowed user to switch between 4G to other networks and vice versa (vertical handover) as compared to the previous generation network that facilitated horizontal handover. Its main purpose is to provide higher speed, quality, and capacity to the users, while improving security and costs of internet, data, voice, and other multimedia services. Mobile web access, video conferencing, cloud computing, gaming, etc., are some of the potential and current applications that the 4G cellular network can provide. In terms of data rates, the 4G cellular network can offer up to 100 Mbps for outdoor environments and 1 Gbps for indoor environments. Apart from this, it also provides a latency of 100 ms while maintaining the same bandwidth of 20 MHz that the previous 3G network provides. In terms of capacity, the 4G networks provide an increased scalable capacity of 50 to 500 bits/s/Hz/km^2 to its base station systems.

3.2.2.3 5G cellular network

5G is the fifth and current generation of cellular network whose intention is to improve the services provided by 4G cellular network. It promises to significantly improve the data rates, provide higher connection density and lower latency [60]. Some of the major features of 5G cellular network includes device-to-device communication, improved battery consumption and overall wireless coverage. As compared to 4G cellular network, the 5G cellular network provides a bandwidth of 1 GHz, data rates up to 10 Gbps and a maximum speed of up to 35.46 Gbps which is around 35 times faster than the speed provided by 4G. Such features of 5G networks will immensely help in the facilitation of applications that require less latency, real-time communication with greater accuracy and precision. Furthermore, it also provides a latency of less than 1 ms and a scalable capacity of 50,000 bits/s/Hz/km^2. In terms of network capacity, this cellular network will be able to support about 65,000 connections at any point in time. As 5G network operates on the radio frequencies of 28 GHz and 60 GHz for increased bandwidth, it therefore utilizes millimetre wave (mm wave) technology. However, the drawback of millimetre waves is that they are highly susceptible to interference and cannot travel long distances without significant attenuation. Therefore, base stations for 5G networks must be employed with a distance of no more than 200 m or 300 m from one other. In terms of size, these base stations will be much smaller than traditional base stations of the preceding generation communication technologies.

4.2.2.4 6G cellular network

6G is the sixth and upcoming generation of cellular network whose intention is to improve the services provided by 5G cellular network. They are mainly developed with the aim to give support to data hungry enabling applications through providing an enhanced connectivity and extended network coverage. In order to provide these connectivity and network coverage demands in a cost-efficient manner, 6G aims to be decentralized in nature and will be designed based upon the integration of various communication networks like marine, aerial, underwater, terrestrial, etc. Such integration will aim to provide high speed internet network to any areas whether it is land, sea, rural, or urban. However, in order to provide such facilities, it is essential to obtain various new infrastructures and architectures with high quality communication services [61].

Table 3-3: Comparison of Different Cellular Communication Networks.

Features	2G	3G	4G	5G	6 G
Latency	500-1000 ms	200 ms	100 ms	10 ms	1 ms
Frequency	1.8 GHz	1.6 – 2 GHz	2 – 8 GHz	3 – 30 GHz	-
Bandwidth	200 KHz	20 MHz	20 MHz	1 GHz	100 GHz
Data Rates	64 Kbps	2 Mbps	1 Gbps	10 Gbps	1 Tbps
Security	64-bit A5	128-bit KASUMI Cipher	128-bit AES, 168-bit DES	D2D and continuous authentication	-
Coverage Range	1-10 km	1-10 km	31 km	200-300 m	< 200 m
Network Capacity	-	-	1000 nodes	65,000 nodes	-
Scalable Capacity	-	10 bits/s/Hz/km^2	50-500 bits/s/Hz/km^2	50,000 bits/s/Hz/km^2	-

3.2.3 Virtual domain layer

The virtual domain layer comprises of the two major sub-layers: data aggregation and modelling and the data analysis and visualization layer. Mentioned below is a brief explanation for both of these sub-layers:

3.2.3.1 Data aggregation and data modelling

The role of the data aggregation in this sub-layer is to aggregate the environmental and operational data that was received by the communication network layer as well as the historical data of the given physical entity, its design specifications, and bill of materials so that it could be used by the data modelling process present in this sub-layer as well as the further upcoming layers. The received data can be of three major types: structured, semi-structured, and unstructured. Figure 3-8 illustrates the different data types that exists in the context of DT. Table 3-4 illustrates a comparison between the different type of the collected DT data that can exist in this sublayer.

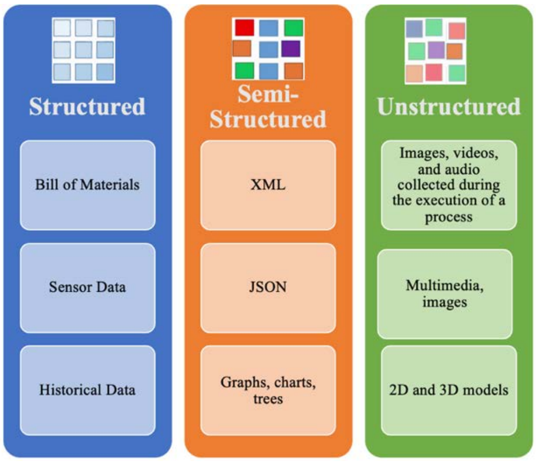


Figure 3-8: Types of Digital Twin Data.

The aggregation process of the data can take place in either the cloud platform or on the enterprise premises itself. The function of the aggregation process is to ingest data that is sent by the earlier physical domain layer as well as the pre-existing data of the physical entity. This pre-existing data must be transferred to a type of infrastructure of databases called data repositories. In general, the two major types of data repositories can be used for DT which is the data lakes and data warehouses. The role of data lakes is to store structure, semi-structure, and unstructured data in its raw form, that is, the form that exists prior to being stored in the repository, tagged with metadata.

Since the data lakes can be easily scalable at an affordable cost. Hence, they can be considered ideal for situations that required rapid storage of large volumes of data. Furthermore, the data stored in the data lakes will only become structured if it is required by the future operations. Data lakes often consists of a combination of numerous technologies that may include cloud storage, databases, and Hadoop Distributed File System. On the other hand, data warehouses can be used to aggregate data over numerous sources like data lakes. However, for data warehouses the data needs to be organized and structured before being stored. Thus, data warehouses compromise of data that are can be readily analysed when accessed. Google Big Query, Azure SQL Data Warehouse, AWS Redshift, Oracle Autonomous Data Warehouse are examples of some of the popular vendor tools for data warehouses that exist in the market.

Table 3-4: Comparison of Structured, Semi Structured and Unstructured Data.

Features	Structured	Semi-Structured	Unstructured
Format	Relational database	HTML, JSON, XML	Binary and character data
Robustness	Very robust	Limited	-
Scalability	Difficult to scale	Simple to scale	Very scalable
Query Performance	Structured query allows complex joining	Anonymous queries are possible	Supports only textual queries
Storage Requirement	Less	Significant	Large

In the case of the DT, the data lakes can be used to rapidly store the real-time operational and environmental data that is collected from the sensors. This data can be structured in accordance with the provided schema. In addition to this, this data can be stored in data warehouses in non-real time. Data warehouses in general can play a vital role to store the previously acquired historical data, bill of materials, and design specifications of the given physical entity. Data stored in both repositories will be used to build dynamic 3D models and for further analysis in the above upcoming layers.

On the other hand, the function of the data modelling sublayer is to design dynamic 3D models of the given physical entity that is being observed. As the physical entity goes through changes, the sensors present in the physical domain layer detects these changes and sends it to the virtual domain layer in which the data aggregation sublayer first stores the received data in the respective data repositories. Later, the data modelling sublayer will read this updated data from the data repositories and then dynamically design the corresponding 3D model. This 3D model takes the updated data into consideration to reflect the present state of the given physical entity. In addition to this, the model can be extended with the physical information as well as the temporal and contextual aspects of the physical entity relative to its operating environment such as the status of the machine, location, process flow, movement, temperature, pressure, and energy consumption. This in turn helps to generate a model that is more dynamic and robust in nature. 3D modelling software can be used to create the model and maps. Various programmatic modelling software options are available in the market that can be chosen. Google Sketchup is an example of one such software where the modelling process can be automated via the use of the Ruby API. Simio Simulation Software is an example of another software that helps to build base 3D models of the given physical entity using design specifications, bill of materials, and historical data provided to it. In addition to this, it can also be used to dynamically update the model based on sensor readings acquired from the physical domain layer.

3.2.3.2 Data analysis and visualization

Once the data is received at the virtual domain layer, the data analysis and visualization layer then can access the data repositories to mine the data and report the result to the management. The purpose of data mining is to examine the condition of the given physical entity to forecast required maintenance or failures that might occur in the near future. Thus, the Virtual Twin of the given physical entity can be used effectively to remotely manage and monitor the physical entity. This management and monitoring of the physical entity can be done by using various data analytics and business intelligence technologies and tools that facilitates the planning of effective maintenance schedules, predict failures and disruptions, etc. With the generated Virtual Twin model, design specifications, and bill of materials data, one can efficiently construct the available raw data in the virtual domain layer into a knowledge base that demonstrates the status and performance of the given physical entity. This will thus help to transform the generated Virtual Twin model into a model that would be driven based on the operational parameters and the subsequent environmental data that are collected by the sensors. In addition to this, the use of advanced machine learning algorithms and data analytic tools can help to integrate real-time streaming sensor data with other operational inputs to generate an operational driven Virtual Twin. This operational Virtual Twin will help to display a dynamic virtual representation of the entire physical entity as well as the processes and operations it executes. Thus, it is important to combine data analysis with DT to obtain the best performance, quality, productivity, and efficiency of the given physical entity. The purpose of the data analytics sublayer is to perform the following two major operations [2]:

- Execute complex operations on the large volumes of data that were collected by the sensors present in the physical domain layer which otherwise would not be able to be performed by the microcontroller.
- Implement suitable machine learning or deep learning algorithms to predict the condition of the physical entity in the near future. First a classifier is built by utilizing supervised machine learning or deep learning algorithms. The purpose of this classifier is to work upon the real-time sensor data, historical data, bill of materials and the design specifications of the given physical entity. Mentioned below are the further elaboration of these techniques:

Execution of Various Data Cleaning and Processing Techniques for Complex Tasks

Various pre-processing and data wrangling methods provide different ways to handle sensor dataset, especially multivariate sensor dataset. Outlier detection, smoothing, data reduction, feature extraction, imputation from the raw sensor data are some examples of different techniques that can be applied on the multivariate sensor dataset. Smoothing on the sensor data is important to eliminate outliers before fitting the data to the respective machine learning model. Whereas discretization can be used to define the sensor values into nominal value based on the interval of threshold value that is pre-defined. Equal width binning, equal frequency, entropy minimization, and

Boolean reasoning are some examples of methods that help to perform discretization on sensory data. The purpose of the discretization method is to provide nominal or categorical data values that are required by most of the data mining algorithms to train its respective models. Attribute reduction techniques like Principal Component Analysis (PCA) can be used to transform the data into an independent set of attributes. In addition to this, it can also be used in selecting those sensor attributes that help to depict the most variation, thereby reducing the complexity in a strongly dimensional dataset while still maintaining the patterns and trends.

Implementation of Supervised Machine Learning Techniques on Sensor Data

Machine learning is a huge field which can be classified into three major classes: supervised, unsupervised and reinforcement learning. The supervised type of machine learning algorithms can be further classified into classification and regression supervised machine learning techniques. In the context of sensor data, both of classification and regression technique types of supervised machine learning techniques can be applied on it to gain a prior knowledge of the input data and its corresponding labels. That is, both the input data and its label can be used to perform various activities like predictive maintenance, quality analysis, efficiency optimization, etc. Table 3-5 furthers shows the major differences between regression and classification techniques. One of the main use cases of classification algorithm is that it can be used to detect the failure of a given particular by examining the relative failure through the labels of failure and no failure in the sensor readings. Artificial neural networks, support vector machines, naïve Bayes, random forest, logistic regression, etc., are some of the common classification algorithms. Since multiple causes can be responsible for potential failure in a given sensor, thus any of these classification techniques can be employed to implement predictive maintenance using multiclass classification. Another type of supervised machine learning technique is regression which can be used to forecast the future sensor readings. In addition to this, the regression techniques can also be used to estimate the Remaining Useful Life (RUL) which is the number of remaining days left before the next failure occurs to a given sensor.

Table 3-5: Comparison between Classification and Regression ML Techniques.

Classification	Regression
Predicts a discrete class label	Predicts a continuous quantity
Data is labelled into one of two or more classes	Requires the prediction of a quantity
Dependent variables are categorical	Dependent variables are numerical
Support vector machines, naïve bayes, nearest neighbours etc., are some examples of classification algorithms	Decision trees, linear regression, random forests, etc., some examples of regression algorithms

Artificial neural networks are another major type of machine learning algorithm that is basically a type of a computational model which is based on the structure as well as the functions of a biological neural network. They are suitable for sensor data that comprises of multiple distinctive parameters over a large time span. This is because the sensors readings are mostly dimensional in nature over a small frequency. In addition to this, they can be used to predict the sensor readings for a lost or dysfunctional sensor, if there exists a strong interdependence between the sensor readings.

3.2.4 Application layer

The purpose of the application layer is to provide personalized services according to the needs of the customers or service providers across various sectors like healthcare, manufacturing, vehicle, retail, etc. However, to provide these services, the application layer first needs to turn the large amount of data transmitted to the virtual domain layer into useful insights by the help of various techniques and businesses intelligence tools that help to record all the activities for a given component or sensor. For instance, in the case of the manufacturing sector, the collected sensor data can be used to analyse to provide suggestions for scheduling predictive maintenances to decrease machine downtime. Furthermore, the collected data can also be converted into useful visuals like dashboards, bar graphs, reports, etc. This visual representation of the collected data can be used to identify and examine problems for a given application. For instance, it can be used in the healthcare sector to draw conclusions related with the health status of a specific patient. The acquired health status can be used to detect the most effective and treatment for the given patient. The generated dashboard charts can also be used to figure out certain Key Performance Indicators (KPIs) that are necessary to execute certain industrial automation decisions. Various alert and notification systems can also be generated based on the processed results obtained in order to help users, organizations and service providers act quickly to address the underlying issues faced by the given physical entity. Moreover, the generated reports can be further analysed to detect any kind of weakness present in the design of the physical system. In case a problem arises in the physical system, necessary actions can therefore be taken to mitigate these problems faced. Therefore, depending upon the application domain, a DT can be thus used to help improve the efficiency, productivity, performance, and quality of a given physical entity. Figure 3-9 illustrates the major application domains that can implement DT to make use of such features.

Manufacturing
Smart Cities
Digital Twins
Healthcare
Automobile

Figure 3-9: Digital Twin Applications.

3.2.5 Security layer

As depicted in Table 3-6, each of the layers of the DT architecture comprises of various vulnerable components. The presence of these components can give rise to numerous security challenges. Attackers can exploit these security challenges to launch various security attacks. For instance, the physical domain layer can be vulnerable to different physical attacks whose goal is to replace, damage, steal the sensors, actuators, or the microcontrollers. Device tampering, firmware attack, reverse engineering etc., are some of the attacks that can occur in this layer. The communication network layer can also be vulnerable to different attacks whose aim is to cause a disruption in the service that the network provides. Man in the Middle attack, DoS attack, replay attack, etc., are some of the examples of the attacks that can occur in this layer. The virtual domain layer on the other hand is also susceptible to various security attacks whose goal is to tamper or steal the data during the data aggregation and modelling phase or during the data analysis and visualization phase. SQL injection, authentication attack, malware injection, etc., are some of the examples of the attacks that can occur in this layer. Finally, the application layer can also be exposed to various security attacks. The goal of the attacker in this layer is to disrupt the personalized services offered to the users based on their requirements. Session hijacking, phishing, malware etc., are some of the examples of the attacks that can occur in this layer. Therefore, to mitigate these attacks, organizations must ensure to implement different security countermeasure techniques that can help to eradicate the security attacks and challenges faced in each of the mentioned layers. To achieve this, it is important for organizations to understand in detail how the different security attacks can occur. Thus, the next chapter focuses in detail on the security aspects of DT.

Table 3-6: Security Concerns at Each DT Layers.

Domain	Vulnerable Components	Security Challenges
Physical Domain Layer	Sensors, actuators, and microcontrollers	Data security and data authenticity
Communication Network Layer	Cellular networks, ethernet, and satellite	Lack of authentication, heterogenous nature of collected data
Virtual Domain Layer	Data lakes, data warehouses, machine learning, deep learning	Data security, guarantee of access control, data breaches
Application Layer	Third party apps and websites	Data security of the employed applications, no universal standards or set policies

Chapter 4. Digital Twin Security

In this chapter, an overview on the importance of security in DT is presented. After which a list of common security attacks and countermeasures that can occur generally in the layers of the DT system is presented.

4.1. Overview

The rise in the implementation of DT has set to transform various industrial and manufacturing processes. According to a recent survey report, the global market size of DT is set to grow from $3.1 billion in 2020 to $48.2 billion by 2026 [62]. This massive rise in the utilization of DT is because of the various benefits it provides such as improved productivity, reduced costs and downtime, optimized maintenance, and monitoring of physical assets, etc. Nevertheless, just like any new technological trends, the security aspect of DT is unfortunately neglected as well, thereby giving rise to a multitude of security vulnerabilities, threats and attacks that can take place. Thus, to have a better understanding and examination of the security aspects of a DT, one must have an idea on how a DT is created in the first place.

The process of creating a DT involves the collection of real-time and operational data that is generated by one or more sensors. Once this data is collected, it is then sent to a cloud-based system where analysis of this data can be performed via various machine learning algorithms. Finally, based on the relayed information, any identified changes are then replicated in the Twin itself. Since a large portion of DT utilizes cloud services for the purpose of storing and processing data. This means that the management and selection of these services needs to be carefully done while bearing in mind the security aspect of such services. Management of such services must ensure that no software applications are executed without being pre-authenticated first. That is, users should be only allowed to access data according to the minimum level of data access they require. In addition to this, all collected data must be encrypted while being stored. Transmission of this data must be performed using secure and encrypted channels. Apart from this, physical security systems and their DT should be also monitored via an automated verification service. The purpose of this service is to ensure that the DT is working as per desired and that no files have been corrupted or accessed by an authorized party without the knowledge of the system administrators. Thus, the presence of security vulnerabilities and threats makes it necessary for organizations to not rush on the adaptation of DT without carefully updating and assessing the latest security protocols. In addition to this, they must also be fully aware of the potential security vulnerabilities and threats that can occur in a traditional DT system. This is because acquiring a secure DT will not only protect the confidential information it stores, but it will also protect the data from corruption that could have a negative impact on the decision-making process.

4.2. General Security Attacks and Countermeasures

As illustrated in chapter 3, a DT can be divided into five major layers: (1) physical domain layer; (2) communication network layer; (3) virtual domain layer; and (4) application layer. Each of these layers uses diverse technologies that brings forward numerous security attacks. Thus, the following section discusses the various possible security attacks and countermeasures that are relevant to the DT system across the five layers. These security attacks and countermeasures were found by an extensive study of multiple research papers.

4.2.1 Physical domain layer

The DT comprises of the physical devices such as the sensors, actuators, and microcontrollers that are vulnerable to various security threats due to the security flaws present in their architectures. The sensors present in the DT can be of various types. Therefore, mentioned below are the security attacks that can occur:

4.2.1.1 Eavesdropping

It is one such attack that can occur in the physical domain layer [63]. Since a typical RFID system compromise of tags and readers that are wirelessly connected and communication without the need of a human intervention. Therefore, there is a high chance that their respective communication medium can be eavesdropped. Typically, the eavesdropping attack is launched when the attacker can get the data that is transmitted between the tag and the reader. This is mainly due to the reason that most of the RFID systems do not compromise of an encryption mechanism during the transmission process because of its memory capacity. Thus, this makes it extremely easy for the attacker to obtain any confidential data from the respective RFID tags.

4.2.1.2 Spoofing

This type of attack happens when a malicious tag pretends to be a valid tag and obtains an unauthorized access [64]. That is, this attack is used to eavesdrop the data coming from the valid tag and copy the captured data to another one This type of attack is concerned with spoofing RFID signals to obtain data stored on an RFID tag. After which the attacker uses the original tag ID to send his own data to appear to be from the original source, which enables the attacker to access the entire system as a legal node.

4.2.1.3 Node tampering

This type of attack targets the sensors node by performing an actual physical damage on it or even replacing the entire node or part of its respective hardware to gain access to confidential information [65].

Apart from the sensors, the physical domain layer also comprises of microcontroller that can also be vulnerable to numerous security threats due to the security flaws present in their architectures. Therefore, mentioned below are some of the most widely known security attacks:

4.2.1.4 Firmware attack

It is an attack that can occur in this layer [66]. This attack arises when the firmware files are not encrypted during transmission. This allows attacker to sniff and modify the traffic that occurs between the microcontroller and the server. Once the transfer of the malicious software is taken place into the targeted microcontroller the attacker would only need the architecture of the microcontroller that can be easily obtained by acquiring its manual [67]. Thus, the attacker would use the flaw that exists in the firmware updates for their advantage and the malicious code would then run once the firmware updates are installed onto the device.

4.2.1.5 Reverse engineering

Through reversing the source code of a given software or firmware used onto the microcontroller, an attacker can thus gain access to the confidential information such as the hardcoded credentials. In addition to this, the attacker can also find any kind of bugs that exits in the code which would help them to plan their attack. Thus, this attack can be performed to inject malicious code onto the microcontroller [68].

4.2.1.6 Malware

An attacker can use a malicious software (malware) to try to infect the given microcontroller. Various types of malwares exist as an option for the attacker to launch this attack [69]. However, a common characteristic that exists between all of them is that they all are unwanted and potentially harmful in nature for the infected microcontroller. An infected microcontroller from a malware can modify the behaviour of the device itself [70].

Mentioned below are the countermeasures that can be implemented to mitigate the attacks faced in the physical domain layer:

4.2.1.7 Physical security design

Most of the devices in this layer can be avoided by designing devices in such a manner that they are physically secure in nature [71]. This includes data acquisition unit design, radio frequency circuits, chip selection, etc. These components are required to be of high quality in nature and they should also not be easily changeable.

4.2.1.8 Authentication

To keep malicious devices away from the DT network authentication of the devices needs to be done before they get into the network. Without the implementation of proper authentication, the device should not be allowed to communicate with the network that can prevent fake data flow into the network.

4.2.1.9 Encryption

Any lightweight encryption could also be implemented to mitigate the possible security attacks that can occur in this layer. The advantage of these encryption algorithms is that it uses less memory, power, energy, as well computing resources to provide the security solution for resource-limited devices.

4.2.2 Communication network layer

The communication network layer of the DT system is also vulnerable to various security threats. These threats arise due to the various security drawbacks present in the communication network. Some of the major security attacks include the following:

4.2.2.1 Denial of Service (DoS)

In this type of attack, the attacker tries to flood the network with large traffic so that the services that are meant to be provided to the intended users are unavailable [72].

4.2.2.2 Man in the Middle (MitM)

In this type of attack, the attacker over the internet tries to intercept the communication that occurs between the two entities [73]. The attacker then disguises themselves as a legitimate node that communicates with the two victim nodes. This way the attacker successfully gains the trust of both the victim nodes and obtain confidential information about them.

4.2.2.3 Eavesdropping

This attack allows the attacker to secretly listen to the communication that can take place between two entities. Thus, by launching this attack the attacker can listen to the conversations when the data is being transferred via unsecure servers, when the device is connected to an unsecure wireless network, when unwanted ports are open, etc., [74].

Mentioned below are the countermeasures that can be implemented to mitigate the attacks faced in the communication network layer:

4.2.2.4 Authentication

The illegal access of any nodes can be avoided by the implementation of appropriate authentication techniques [75].

4.2.2.5 Encryption

Common encryption algorithms like AES, DES etc., can also be implemented to mitigate the attacks faced in this layer.

4.2.3 Virtual domain layer

The virtual domain layer comprises of the two sub-layers: The data aggregation and modelling and the data analysis and visualization. Many security threats can occur in both sub-layers.

Therefore, in terms of the data aggregation and modelling sublayer, mentioned below are some of the major security attacks:

4.2.3.1 SQL injection

This type of attack is used by the attacker to gain administrative access to the databases via targeting the vulnerabilities that exist in the network [76]. To perform this attack, the attacker needs to input SQL commands in web forms that would execute fraudulent commands which would help to retrieve sensitive information. If the given database is not fully secured, then it will run these fraudulent commands without any issue and thus would provide the attacker access to the confidential information upon request.

4.2.3.2 Data modification

In this type of attack, any entity who has access to the sensitive information storage technologies can modify the data for their personal benefit or financial gain. During this attack, the attacker would try to manipulate the data and extract the sensitive information from inside.

4.2.3.3 Unauthorized attack

In this type of attack, an unauthorized attacker tries to infiltrate the system and also prevent the legitimate users to access into the system itself [77]. In addition to this, the attacker also can try to delete the confidential information.

In terms of the data analysis and visualization sublayer, mentioned below are some of the major security attacks that can occur:

4.2.3.4 Password attack

This type of attack can be launched by various means like password guessing and password re-usage [78]. Password guessing as the name implies is a type of attack in which the attacker tries to input common password combinations until they successfully find a password match. Whereas password re-usage is another type of attack in which the attacker knows the password already, however they try to use this same password to access multiple other accounts.

4.2.3.5 Malware

Malware is another type of attack that could occur in the cloud [79]. In this attack, the attacker tries to alter data, obtain control, or execute malicious code through injecting malicious service instance or virtual machine into the cloud [80]. For instance, the attacker would copy and upload the service instance of the victim, but malicious instance responds to the request when some service requests the instance of the victim. As a result, the attacker would thus obtain the confidential data.

4.2.3.6 Authentication attack

Another form of attack is called the authentication attack [81]. This attack can be used by the attacker as some of the cloud services still use some sort of a single factor authentication process and a simple username and password requirements or compromise of weak password. The attacker can therefore utilize this vulnerability to their advantage when they want to try to disrupt the services or steal the confidential information.

To mitigate the security attacks faced in this layer, mentioned below are some of the countermeasures that can be employed:

4.2.3.7 Data encryption

One such example is homomorphic encryption. In this type of encryption, the ciphertext is allowed to be computed immediately without being decryption. However, it requires a high computation although it assures data security [82].

4.2.3.8 Fragmentation Redundancy Scattering (FRS)

In this technique, the important data onto the cloud is separated and allocated to various fragments of storage in servers. Since the fragment has no useful data details by themselves alone, the risk for data theft is therefore highly minimized [83].

4.2.4 Application layer

The application layer of the DT system comprises of tailored services that are presented to the user according to their desired preferences. The major security threats to the application layer target these personalized user services. Some of these attacks include the following:

4.2.4.1 Session hijacking

In this attack, the attacker tries to create a legitimate temporary session that would be then used between the server and the application user for harmful purposes [84]. To generate such a session, the attacker would first use some form of an authentic cookie from the computer of the user to be able to connect to the server. After performing this action, a valid connection is then established. This connection would then be used to access any information that is transmitted between the server and the user.

4.2.4.2 Phishing

In this attack any sort of sensitive information can be obtained by the attacker as they would masquerade themselves as a legitimate entity [85]. Often fake links or forged emails that look like some sort of a legitimate websites are used to deceive the user in believing that they from a trustworthy entity to gather the confidential information about them.

4.2.4.3 Malware

In this type of attack, the attacker would create some sort of a malicious software that would then get installed on the device of the victim without them knowing about it. This attack would be then used to gain access to the confidential information or to damage the device of the victim itself [86].

Mentioned below are the countermeasures that can be implemented to mitigate the attacks faced in the application layer:

4.2.4.4 Anti-virus, anti-spyware, and anti-adware

Various software like anti-virus, anti-spyware, and anti-adware needs to be installed in the systems of the user to ensure the confidentiality, reliability, and integrity of the DT network [87].

4.2.4.5 Risk assessment

Another way the application layer can be secure is by the implementation of risk assessment technique. The purpose of this technique is to continuously detect threats which occurs in the system, apply relevant patches, and update the firmware of the system so that the overall security of the system can be enhanced further [88].

4.3. Proposed Digital Twin Security Model

As observed from the above subsection, DT can be vulnerable to numerous security attacks. Thus, through the analysis of these security attacks and its relevant countermeasures a general proposed DT security model was presented as illustrated in Figure 4-1. Based on this, a multi-layer secure DT architecture was proposed as illustrated in Figure 4-2.

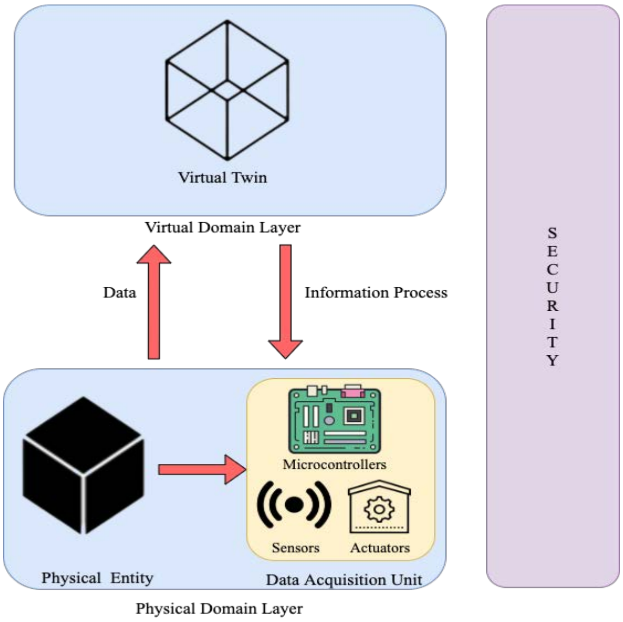


Figure 4-1: Proposed Digital Twin Security Conceptual Model.

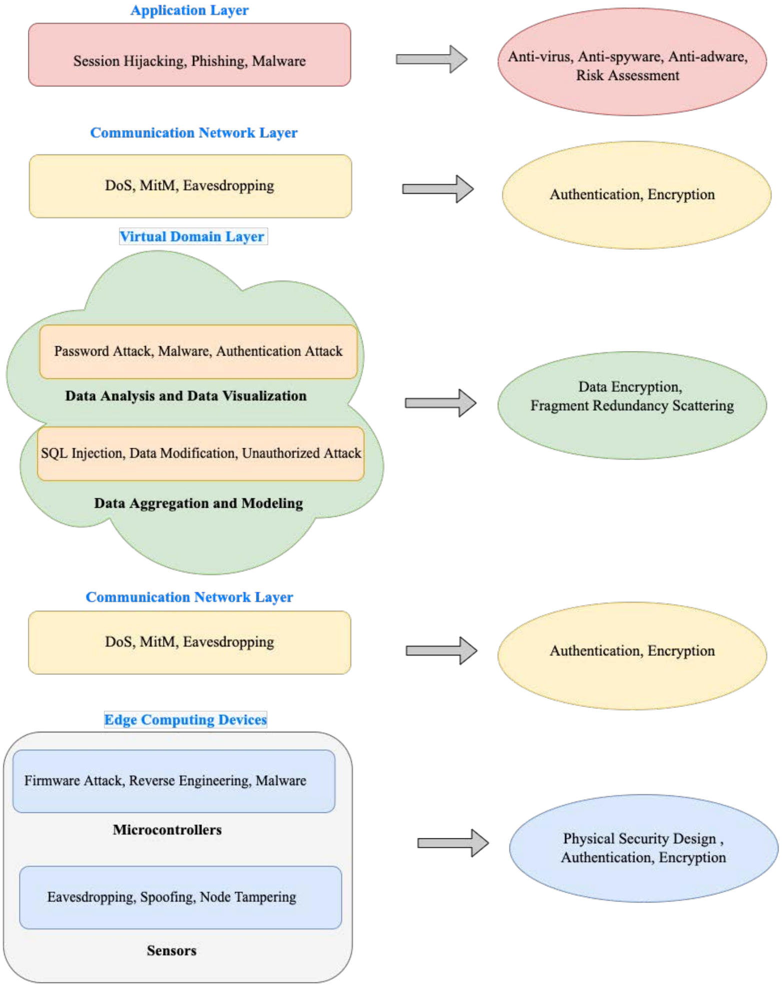

Application Layer
Session Hijacking, Phishing, Malware
Anti-virus, Anti-spyware, Anti-adware, Risk Assessment
Communication Network Layer
DoS, MitM, Eavesdropping
Authentication, Encryption
Virtual Domain Layer
Password Attack, Malware, Authentication Attack
Data Analysis and Data Visualization
SQL Injection, Data Modification, Unauthorized Attack
Data Aggregation and Modeling
Data Encryption, Fragment Redundancy Scattering
Communication Network Layer
DoS, MitM, Eavesdropping
Authentication, Encryption
Edge Computing Devices
Firmware Attack, Reverse Engineering, Malware
Microcontrollers
Eavesdropping, Spoofing, Node Tampering
Sensors
Physical Security Design , Authentication, Encryption

Figure 4-2: Proposed Multi-Layer Secure Digital Twin Architecture.

Chapter 5. Digital Twin Security Modelling

This chapter talks in brief about the proposed method used and its relevant background.

5.1. Model Checking Background

To implement the proposed multi-layer secure DT architecture mentioned in the earlier section, a formal modelling technique was used. Although, several formal verification techniques are extensively used to implement and verify a given system design across various domains. Model checking and theorem proving are some of the well-known formal verification techniques that has also been used as a validation technique. Therefore, in this case, an approach that uses Probabilistic Model Checking (PMC) to validate the system properties was chosen. This approach is different than other existing approach such as machine learning. The reason for this is because PMC approach relies on an existing model where it can check certain properties of the system represented in temporal logic, whereas the machine learning approach learns the model and then predicts the outcome on a new dataset.

The models chosen for the PMC model are coded using PRISM language whereas the property that needs to be investigated is written in property specification language that incorporates various temporal logics. Some examples of the common models used in PRISM includes the following:

1. Markov Decision Process (MDP)
2. Discrete-time Markov Chain (DTMC)
3. Continuous-time Markov Chain (CTMC)
4. Probabilistic Automata (PA)
5. Probabilistic timed Automata (PTA)

Mentioned below are the following two models used to validate the system properties:

5.1.1 Markov Decision Process (MDP)

A typical MDP is formally represented by a 4-tuple $\mathrm{M} = (\mathrm{S}, \mathrm{s}', \mathrm{A}, \gamma)$, where S is a finite number of states, s' is the initial state where $s' \in S, A$ represents the set of all possible actions, and finally γ refers to probability distribution for every state $s \in S$, $a \in Act(s)$ such that $\sum_{s' \in S} \gamma\,(\mathrm{s}, \mathrm{a}, \mathrm{s}') = 1$. Where $Act(s)$ refers to the set of enabled actions at the state s . To demonstrate the working of the MDP, a simple example of a MDP model is demonstrated in Figure 5-1. This model comprises of the following parameters: $S = \{s_0, s_1, s_2\}, s' = s_0$, $A = \{a_0, a_1, a_2\}$, $\gamma(\mathrm{s_0, a_0, s_1}) = \gamma\,(\mathrm{s_0}, a_1, s_2) = 0.5$, $\gamma\,(\mathrm{s_1}, a_2, s_1) = 0.8$, $\gamma\,(\mathrm{s_1}, a_2, s_2) = 0.2$, $\gamma\,(\mathrm{s_2}, a_0, s_0) = 0.4$, $\gamma\,(\mathrm{s_2}, a_2, s_2) = 0.6$. For this case, the enabled actions are as follows: $Act(s_0) = (a_0, a_1), Act\,(s_1) = (a_2), Act(s_2) = (a_0, a_2)$. As seen from the set A, that represents all possible actions in the given MDP model in which the model non-deterministically choses the best action (to maximize or minimize reward) in order to transition to the next state. This behaviour coincides with the way a skilful attacker works, in which case the attacker uses its intelligence and knowledge to choose the best component to launch an attack amongst the available number of components present. Therefore, to model the behaviour of the skilful attacker, the MDP model was chosen.

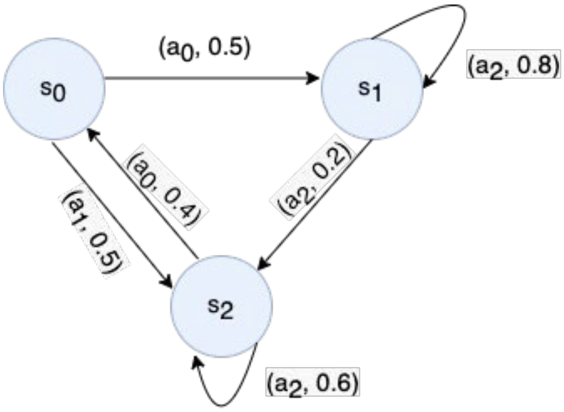


Figure 5-1: Example of a MDP Model.

5.1.2 Discrete-time Markov Chain (DTMC)

Apart from the MDP model, a brief overview of the Discrete-time Markov Chain (DTMC) model is presented in this subsection [89]. Formally, a DTMC can be described as 4-tuple $\mathrm{D} = (\mathrm{S}, \mathrm{s}', \mathrm{P}, \mathrm{L})$. Here, S represents the finite set of states, s' is the initial state where $s' \in S$, P is the transition probability matrix and finally F is the functioning label states that comprises of atomic propositions. That is, $F : S \rightarrow 2^{AP}$. As demonstrated in Figure 5-2, a typical DTMC comprises of 4 states $S = \{\, s_0, s_1, s_2, s_3\}$, here the initial state $s' \in S$. The transition probability for this case is the following:

$$P = \begin{bmatrix} 0 & 0.7 & 0.3 & 0 \\ 0 & 0 & 0.1 & 0.9 \\ 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

Moreover, the set of atomic propositions in this case is $\{x, y\}$ and for the function L the labels is $s_0$with x and s_3 with y respectively. However, in this case, unlike the MDP model, the DTMC model does not compromise of the set A which represents all possible actions. That is, the DTMC model does not have any choice between the actions as only one action is available at any given state. Therefore, the DTMC model does not differentiate between minimum or maximum reward. Thus, in this case, the model deterministically chooses the next best state. Since this behaviour coincides with the way a naïve attacker behaves in which the attacker does not choose the best component based upon a given situation. Therefore, to model the behaviour of the naive attacker, the DTMC model was chosen.

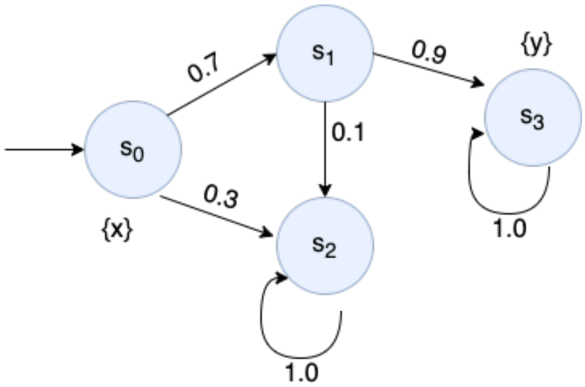


Figure 5-2: Example of a DTMC Model.

5.1.3 Linear Temporal Logic (LTL)

To analyse a probabilistic model which has been specified and constructed in PRISM, it is necessary to identify one or more properties of the model which can be evaluated by the tool. The required specification can be given in the form of linear temporal logic (LTL) which is nothing but an extension of temporal logic that can be implemented to verify certain properties for a given computer system. PRISM is used to either check if the system satisfies the given specifications under a strategy, or to synthesize a strategy that meets some specifications. Mentioned below is one such example of the syntax used in the LTL formula [90] as demonstrated in Equation 1:

$$\Psi ::= \top \mid \Psi \mid \neg\Psi \mid \Psi 1 \wedge \Psi 2 \mid \bigcirc \Psi \mid \Psi 1 \cup \Psi 2 \qquad (1)$$

5.1.4 Probabilistic Symbolic Model Checker (PRISM)

PRISM is one of the widely used probabilistic model checker that has been successfully applied to model and assess various reactive systems that exhibit random as well as non-deterministic behaviour. A variety of systems have been analysed with the help of PRISM such as security protocols, randomized distributed algorithms, network protocols, and transportation systems. A system can be broken into modules that interact with each other thereby resulting in system evolution.

In order to construct and analyse a model with PRISM, it must be specified in the PRISM language, a simple, state-based language, based on the Reactive Modules formalism of Alur and Henzinger [91]. This is used for all of the types of models that PRISM supports. The fundamental components of the PRISM language are modules and variables. A model is composed of several modules which can interact with each other. A module contains several local variables. The values of these variables at any given time constitute the state of the module. The global state of the whole model is determined by the local state of all modules. The behaviour of each module is described by a set of commands which can take the following form:

[action] guard -> prob_1 : update_1 + } {... + prob_n : update_n;

On the other hand, the state variables are declared in the following manner:

x*: bool init false;*

The above statement means to declare a boolean state variable x is initialized to false.

A module in PRISM language starts by using the following command:

module

xyz ...

endmodule

Furthermore, any model can also include certain rewards added to it to provide more information regarding the system constructed. To formulate rewards in the PRISM language, the same syntax to create a module can be used, except the keyword "rewards" would be used instead of "module". Mentioned below is an example of a simple reward that can be used:

rewards

x=0 : 100;

*x>0 & x<10 : 2*x;*

x=10 : 100;

endrewards

5.2. Proposed Methodology

In this section, an elaboration on the proposed approach to model the security in the DT system is presented. The proposed model comprises of two tiers of defence to capture the security of DT for the implementation of various realistic scenario. Furthermore, the model is also independent of individual defence mechanisms that are used at each tier. Any kind of existing security mechanism or services can be employed to mitigate security attacks. In the first tier of defence, individual component-based security mechanism such as encryption and authentication are considered. The defences present at this tier will become weak when vulnerabilities are discovered at the component level. Apart from the defence available at the first tier, there are also exists independent defence mechanisms present at the second defence tier whose role is to help monitor the services and then execute the required actions. Anti-malware software, firewall and intrusion detection system are some examples of the defences that can exist at the second-tier defence.

5.2.1 Overall system model

Let $S_{DT} = (M_l, M_o)$ be a tuple that represents the proposed security system for DT where M_l represents layered security model, and M_o represents the attack objective. M_l is a tuple $\{l_i\}$ where $l_i \epsilon \{l_p, l_c, l_v, l_a\}$ represents physical domain, communication, virtual domain, or application layers. A set of components $C = \{c_s, c_a, c_m \dots\}$ is considered in the system such as sensors, actuators, microcontrollers, etc., in which different kinds of security attacks can be launched on the system. Furthermore, $O = \{o_1, o_2, \dots o_m\}$ is the set of attack objectives considered such that $\forall\ o_i\ \epsilon\ O, C_i \subseteq C$ is a set of components where $c_j\ \epsilon\ C_i$ is vulnerable to o_i and $\cup_{o_i \in O} C_i \subseteq C$. Corresponding to each attack objective $o_i \epsilon O$ and the relevant components $c_j \epsilon C_i$ vulnerable to o_i, there is a tuple $p^{ij} = \{p_o^{ij}, p_d^{ij}\}$ that captures the probability of attack objective and defence being successful respectively. A matrix $P\ [m \times n]$ captures all such tuples with each row $1 \leq i \leq m$ for $o_i \epsilon O$ and column $1 \leq j \leq n$ for $c_j \epsilon C$ such that $\forall\ i\ P[i,j] = \{0,1\}$ where $c_j \notin C_i$.

Depending upon the adversary's objective, there could be multiple methods to achieve it. Each method involves compromising a set of components. The attacker chooses an attack and its corresponding component to launch an objective. After an attack objective is successful with probability p_o^{ij}, the first-tier defence for the system automatically responds to it with the below probabilities:

$$T1_{DEF} = \begin{cases} detects\ with\ (1 - \ p_o^{ij}) \\ does\ not\ detect\ with\ p_o^{ij} \end{cases} \tag{2}$$

In the case the attacker fails to launch an attack to the given component, then the attacker resorts to attacking another component and continues to launch an objective on the given system until the attack objective is met. However, if the attacker becomes successful in launching an attack at the first level, the second-tier defence for the system takes charge and tries to stop the attack with a probability of p_d^{ij}.

$$T2_{DEF} = \begin{cases} detects\ with\ \ p_d^{ij} \\ does\ not\ detect\ with\ (1 - p_d^{ij}) \end{cases} \tag{3}$$

The attack on a given component j is said to be successful if the component cannot defend itself with probability $\left(1 - \ p_d^{ij}\right)$. For such a case, the component is then marked as compromised and the attacker might have to continue its objective on other relevant components till the desired objective is met. Conversely, if the second tier of defence succeeds (with probability p_d^{ij}) in mitigating the objective then the system creates an alert to increase the defence level for the system by a factor def_f (< 1) upto a maximum value $Level_{def}$.

$$Level_{def}^{ij} = \begin{cases} \left\lceil \dfrac{\log\left(\frac{1}{p_d^{ij}}\right)}{\log\,(1+def_f)} \right\rceil & , p_d^{ij} \; > \; 0 \\ 1 & , otherwise \end{cases} \tag{4}$$

The security system operates in parallel, where its job is to constantly monitor the system logs using advanced data visualization and analytics methods to identify any potential threats. The system quickly responds to the recognized threats through proactive measures that helps to significantly reduce the system compromise as shown below:

$$p_d^{ij} = \begin{cases} p_d^{ij} \,.\, (1+def_f)^{l^{ij}} & , l^{ij} \le Level_{def}^{ij} \\ 1 & , otherwise \end{cases} \tag{5}$$

Corresponding to each type of attack and the related components, there is cost $cost_{ij}$ associated to each attempt of attack. Attacker's cost can be given as:

$$Cost_o = \sum_{o_i \in O} \sum_{c_i \in C_i} n_{ij} \times cost_{ij} \tag{6}$$

Where n_{ij} is the number of attack attempt on component c_j for attack objective o_i. The attacker stops if any of the following conditions hold:

1. Any given attack is successful, and the system is compromised against the given attack objective.
2. Attacker tries a wide variety of attacks till it exhausts the provided resources: $Cost_o \geq T_{max}^{cost}$. Here T_{max}^{cost} is the maximum threshold value for the attacker's cost.
3. Attacker has no more time left for time-bound attacks.

Given a set of objectives O and a function $\forall_i \in O, comp(i) = C_i$ which returns the set of components corresponding to each type of objective, the system is said to be compromised if the following formula is true:

$$\emptyset = \bigvee_{i \in O} \emptyset_i \tag{7}$$

Such that $\forall\, o_i \in O, k = |C_i|$ and $\forall_j \in \{1,2, \ldots k\}$ φ_j is true when $c_j \in C_i$ is compromised with probability $(1 - p_d^{ij})$.

$$\emptyset_i = \varphi_1 o\, \varphi_2 o\, \varphi_3 o \ldots o\, \varphi_k \tag{8}$$

Here $o \in \{\vee, \wedge\}$ represents either conjunction or disjunction operator and it is chosen such that $\emptyset_i$ will become true.

5.2.2 Attack-defence model

The construction of the attack-defence model mimics the behaviour of the attacker which can be as mentioned earlier, either a naive attacker or a skilful attacker. A naive attacker is considered to have basic level of knowledge and skills required to launch an attack. Whereas a skillful attacker is an individual who has enough experience and knowledge and skills that are required to execute an attack. Figure 5-3 depicts the attack-defence model used.

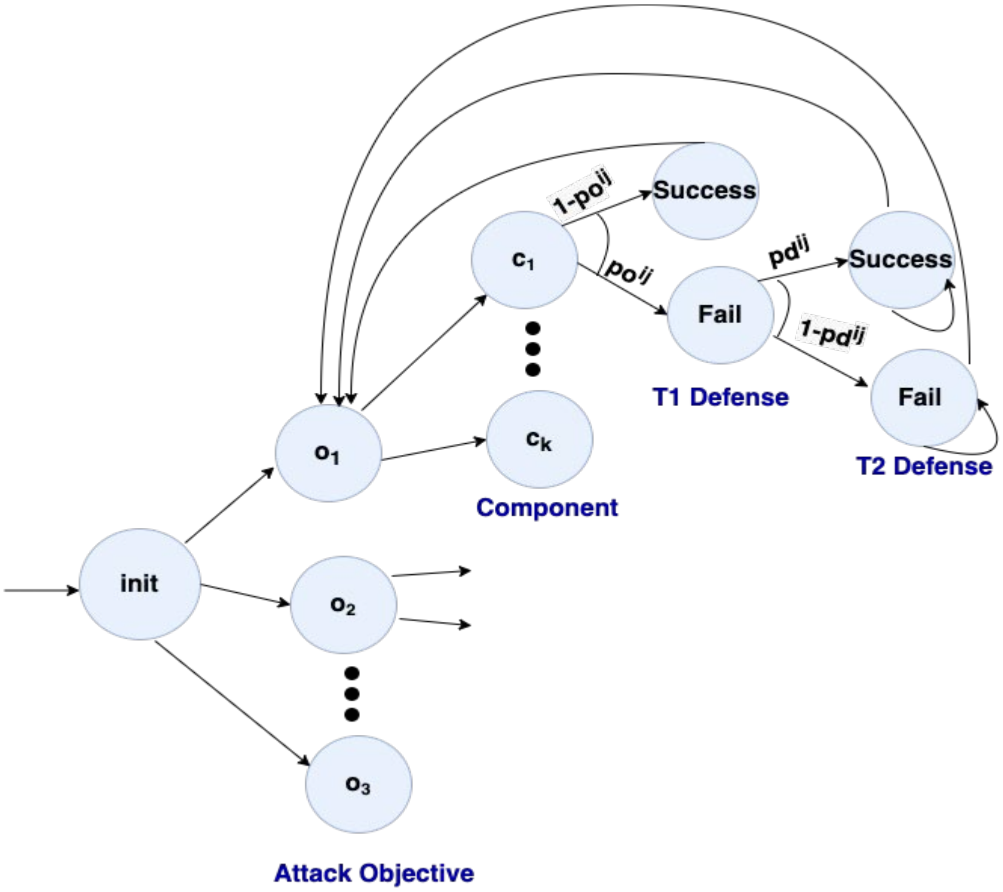
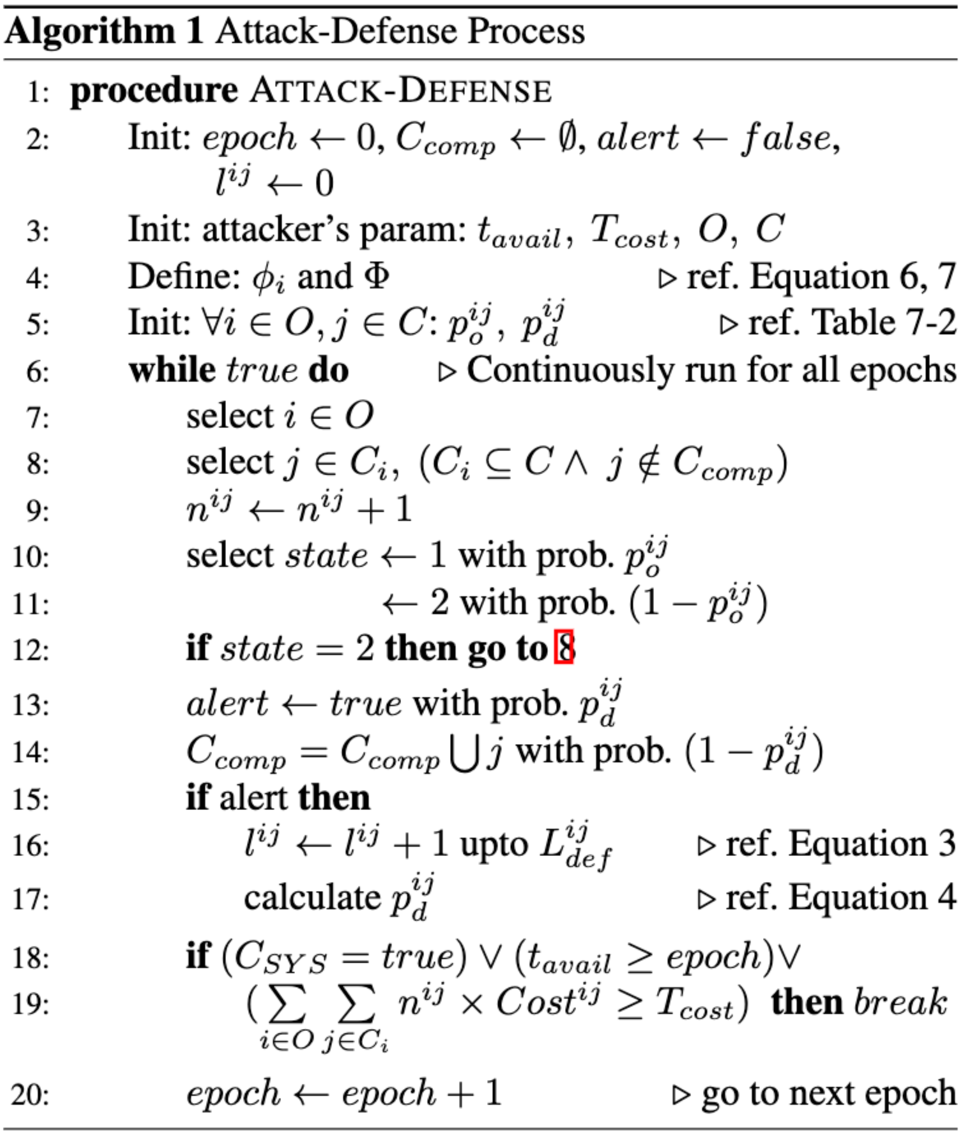


Figure 5-3: Attack-defence Model.

The DTMC (naïve attacker) or MDP (skillful attacker) model starts with the attacker initially choosing a specific kind of attack i from the given list of attack objective O. After this, the attacker chooses a component $c^{ij} \in C_i$ which is modeled non-deterministically or deterministically (depending upon the type of attacker). Once the attacker successfully comprises the chosen component with the given probability p_o^{ij}, it then enters to the next state where the second tier defence activates and tries to mitigate the attack with probability p_d^{ij}. If the second tier defence fails to protect the component against the attack, then the status of the component changes to compromised. The overall working of the proposed attack-defence model is illustrated in Algorithm 1.



Algorithm 1 Attack-Defense Process

```
1: procedure ATTACK-DEFENSE
2:     Init: $epoch \leftarrow 0$, $C_{comp} \leftarrow \emptyset$, $alert \leftarrow false$,
           $l^{ij} \leftarrow 0$
3:     Init: attacker's param: $t_{avail}$, $T_{cost}$, $O$, $C$
4:     Define: $\phi_i$ and $\Phi$                         ▷ ref. Equation 6, 7
5:     Init: $\forall i \in O, j \in C$: $p_o^{ij}$, $p_d^{ij}$          ▷ ref. Table 7-2
6:     while $true$ do          ▷ Continuously run for all epochs
7:         select $i \in O$
8:         select $j \in C_i$, $(C_i \subseteq C \wedge j \notin C_{comp})$
9:         $n^{ij} \leftarrow n^{ij} + 1$
10:        select $state \leftarrow 1$ with prob. $p_o^{ij}$
11:                      $\leftarrow 2$ with prob. $(1 - p_o^{ij})$
12:        if $state = 2$ then go to 8
13:        $alert \leftarrow true$ with prob. $p_d^{ij}$
14:        $C_{comp} = C_{comp} \bigcup j$ with prob. $(1 - p_d^{ij})$
15:        if alert then
16:            $l^{ij} \leftarrow l^{ij} + 1$ upto $L_{def}^{ij}$          ▷ ref. Equation 3
17:            calculate $p_d^{ij}$                      ▷ ref. Equation 4
18:        if $(C_{SYS} = true) \vee (t_{avail} \geq epoch) \vee$
19:            $(\sum_{i \in O} \sum_{j \in C_i} n^{ij} \times Cost^{ij} \geq T_{cost})$ then $break$
20:        $epoch \leftarrow epoch + 1$                  ▷ go to next epoch
```

After the implementation of the model in PRISM is completed, the given system specifications can be generated using LTL formula which is then verified against the proposed model. Mentioned below are the given LTL formula that have been used to check against some of the system properties:

- {P_{max}=? [true U<= Time objectiveSuccess]} here P_{max} stands for the maximum probability that can be achieved by an attacker through the successful execution of an attack objective successfully within the given time constraint. A low P_{max} value means that it is difficult for the attacker to launch an attack objective. Whereas a high P_{max} value means that it is easier to launch an attack objective.
- R {"Attack_Cost"} min=? [C<=Time] where C stands for cumulative and Attack_Cost is a variable used in PRISM to increment the cost value whenever the given attack objective is executed.

Furthermore, the probability of success for a given objective is given by Equation 9.

$$p_o = \alpha^{c_{level}} . p_{\max}^{o} \tag{9}$$

Here $p_{\max}^{o}$ refers to the maximum probability of attack objective success and $\alpha^{c_{level}}$ represents how easy it is to execute an attack on a given component for a specific attack.

Apart from this, for the second-tier defence, three categories of defences: low, medium, and high level were employed as shown below in Equation 10:

$$p_d = \begin{cases} \beta_l . p_{\max}^{d} & low\ level\ defence \\ \beta_m . p_{\max}^{d} & medium\ level\ defence \\ \beta_h . p_{\max}^{d} & high\ level\ defence \end{cases} \tag{10}$$

Here, β_l , β_m, β_h, represents the overall security level values for the respective defence profile levels. And $p_{\max}^{d}$ represents the maximum defence mechanism that exist for the least secure component.

Chapter 6. Use Case: Digital Twin in Healthcare

Although the proposed security DT architecture can be used for any use cases in DT, in this case, it is used as an application for the healthcare sector. Therefore, in this chapter, a case study for the use of DT in healthcare was focused where the different security attacks that can occur in the DT for the healthcare sector in each of the DT layers were presented.

6.1. Overview

The role of DT in the healthcare industry is to generate a virtual replica for a given physical service or object it represents, thereby providing various facilities like evaluation and monitoring without being in near distance with the given physical service or object they aim to represent. In addition to this, the generated Virtual Twin can also help in providing an environment that helps to examine the impact between the performance differences that can occur when any changes would have been executed to the given physical entity. By executing different changes to the physical entity, several possible future problems can be predicted along with the time it might require implementing the necessary procedures or the changes. Additionally, the most efficient and optimal solution can also be chosen accordingly to which solution would lead to risk reduction that can be particularly very crucial for the healthcare sector. The working of DT in healthcare can be broadly divided into the architectural layer it comprises of. That is, initially the process would start with the physical domain layer in which the physical world could comprises of a patient, healthcare object, or service. This physical entity would be then attached to various sensors according to its respective applications. The role of the sensors is to collect the necessary operational and environmental data from the given environment and feed it to the microcontroller. The purpose of the microcontroller is to further process the collected data and send it to the upcoming layers. Apart from this, the data processed by the DT can also be sent back to the actuators to execute some local decisions if required. After the data has been processed by the DT, it will be transmitted to the communication network layer. The communication network layer plays a vital role in acting like a bridge between the physical domain layer and the virtual domain layer. The virtual domain layer plays the role of executing two major roles that can be divided into the form of two sub-layers. The first sub-layer is the data aggregation and modelling sub-layer where the patient vital information gets stored in the data warehouses or in data lakes in the form of EHR/ EMR/ PHR. After this the data gets collected to be further transmitted to the second sub-layer which is the data analysis and visualization sublayer. Here the data gets analysed through the means of different data analytics and business intelligence technologies that help to generate the Virtual Twin that could be the human body, device, or an entire hospital itself. This Virtual Twin is then accessed by the healthcare professional through the means of different devices, websites, or applications.

In general, the role of a DT in the healthcare sector can be broadly classified into three categories: hospital design, hospital management and patient care. Several companies have developed DT in these categories. For instance, a DT for hospital design and management has been developed by GE Healthcare [92]. Dassault Systèmes is another company that has developed a DT for patient care where they claimed to be the first company that has successfully modelled a realistic human heart that provides a virtual representation of all functions of the heart like mechanics, electricity, blood flow, etc., [93]. Famous company like Philips have also benefited from the potential benefit that DT in the healthcare provides when it comes to designing a model of the devices or humans [94]. All the products and services that these companies demonstrate a good example of how one of the biggest advantages of DT is that it can help in testing whether a particular decision would work in the simulated real-time environment or not. Based on how this environment behaves, the DT can give some form of feedback of how efficient the decision was. This decision can be done through the analysis of the simulated environment. A good example of such a situation is when a doctor wants to treat a particular patient. To do so the doctor would run different tests regarding the various treatment on the Virtual Twin of the patient first. The Virtual Twin will then provide feedback on the most optimal treatment for the patient. Thus, a bidirectional communication is established between the healthcare professional (doctor in this case) and the analysis of the data that helps to generate the real-time Virtual Twin. Thus, the adaption of a DT will lead to the generation of a behavioural model that facilitates the experimentation of new healthcare treatments with a in real-time simulated version of the "original" patients before they have been executed.

6.2. DT Security in Healthcare

DT have been recently playing an active role in the enhancement of the healthcare facilities. By creating a Virtual Twin of an operational strategies, hospital, staffing, etc., appropriate decisions can be implemented based on what actions are required to be executed. These generated Virtual Twins can help to solve a numerous problem. In terms of hospital management, they can help to solve problems such as staff schedules, bed shortages, operations rooms, etc. Whereas in the context of patient care they can help to optimize the patient treatment, cost, and effectiveness. Apart from this, a Virtual Twin of an entire hospital can also be generated to create a safe environment which would test the performance impact when any changes occur over it.

Overall, the implementation of DT in healthcare can help to enable efficient strategic decisions, model individual human behaviours based on their individual genomic makeup, physiological characteristics, and lifestyle to suggest personalized medicine. However, the development of a Virtual Twin for a human body would compromise of a much more advanced process that includes sensors whose role is to efficiently provide data for the generation of a Virtual Twin. Nevertheless, before taking advantage of the benefit of implementing a DT, it is important to first consider the security aspect of DT in the healthcare sector. This is because the negligence of security in the healthcare sector could lead to the following possible consequences:

6.2.1 Risk for human life

Any attack that can occur in the Twin architecture for the in-patient monitoring system may directly affect the functionality of the system thereby risking the life of the patient in great danger [95].

6.2.2 Lack of data privacy

Execution of security attacks done by the attacker might also lead to the exposure in the data of the patient that would violate the common data privacy law in the healthcare sector [96].

6.2.3 Loss of reputation

The execution of the attacks done by the attacker would lead to the violations of several healthcare policies and privacy rules that would lead to a loss in integrity which would thereby cause a negative impact on their reputation [97].

6.2.4 Monetary loss

Numerous damages can be caused by the execution of security attacks on the DT architecture. The need to recover from these damages would lead to the extra expense that can have an impact on the overall financial budge of the respective healthcare organization [98].

Therefore, to mitigate these consequences faced, it is important to study its cause (security attacks) so that appropriate countermeasures can be implemented. Thus, the next subsection focuses on the major security attacks faced in terms of the use of DT in the healthcare industry.

6.3. Security Attacks

As seen from the previous chapter, DT can be vulnerable to various types of security attacks. This section focuses on a specific application of DT in healthcare. Since there exist numerous applications of DT in healthcare. In this case study the focus was narrowed down to the use of DT in-patient monitoring system. Figure 6-1 illustrates a typical DT architecture in the use of the in-patient monitoring system. Whereas Figure 6-2 provides a sequence diagram for the in-patient monitoring system. As seen from the Figure 6-1, the working of the healthcare monitoring system can be divided into the following three main categories:

6.3.1 Communication between sensors and microcontrollers

The DT comprises of the microcontroller. Since the data initially flows from the information collected by the sensors attached to the body of the patient to the microcontroller as illustrated in Figure 6-3. Therefore, mentioned below are the following possible security attacks that can occur during the data flows that takes place between the sensors and microcontrollers:

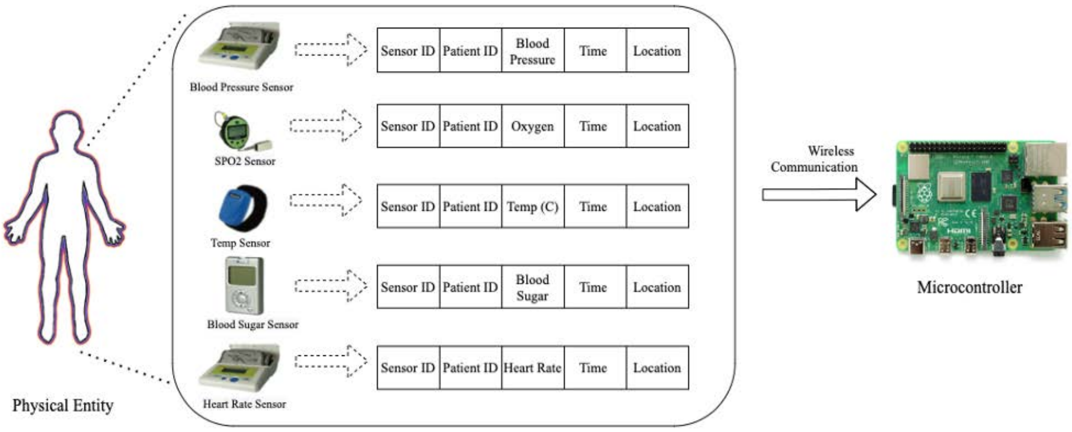

Blood Pressure Sensor
Sensor ID | Patient ID | Blood Pressure | Time | Location
SPO2 Sensor
Sensor ID | Patient ID | Oxygen | Time | Location
Temp Sensor
Sensor ID | Patient ID | Temp (C) | Time | Location
Blood Sugar Sensor
Sensor ID | Patient ID | Blood Sugar | Time | Location
Heart Rate Sensor
Sensor ID | Patient ID | Heart Rate | Time | Location
Physical Entity
Wireless Communication
Microcontroller

Figure 6-1: Data Flow from Sensors and Microcontrollers.

6.3.1.1 Man in the Middle Attack (MitM)

This attack will utilize ARP cache poisoning to intercept packets that is transmitted between two entities. The attacker will act as a Man in the Middle between the sensors and microcontrollers. Thus, the data packets that will be sent to the receiving entity will be first intercepted by the attacking machine and then forwarded to the designated receiver. Figure 6-4 illustrates a visual representation of this attack. In terms of the in-patient monitoring system, the attacker will use this attack to sniff the patient data that gets transmitted from the sensor to the microcontroller.

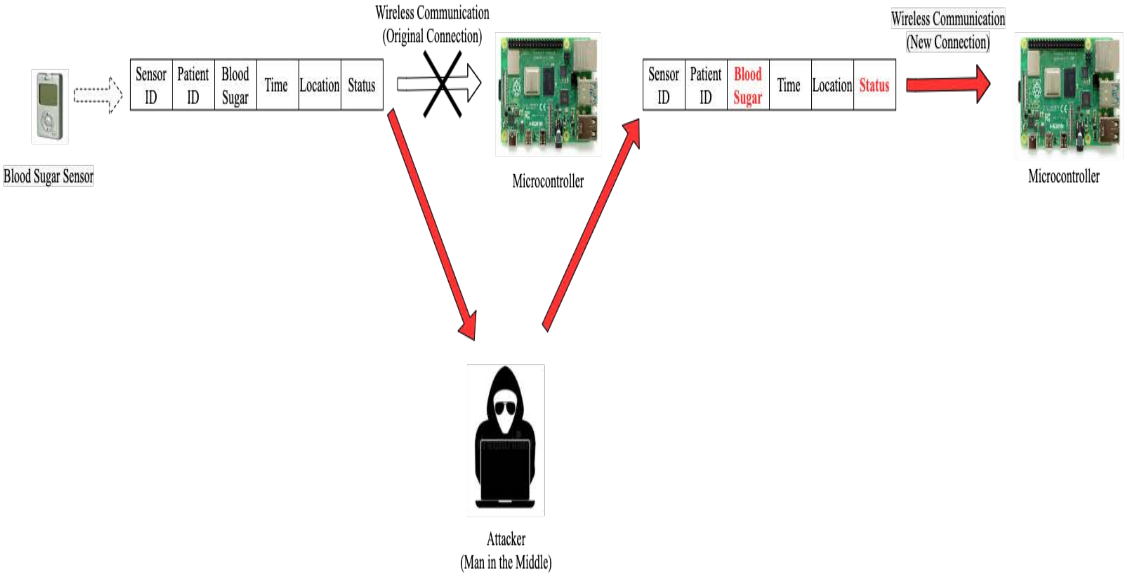

Blood Sugar Sensor
Sensor ID | Patient ID | Blood Sugar | Time | Location | Status
Wireless Communication (Original Connection)
Microcontroller
Sensor ID | Patient ID | Blood Sugar | Time | Location | Status
Wireless Communication (New Connection)
Microcontroller
Attacker (Man in the Middle)

Figure 6-2: Man-in-the-Middle Attack.

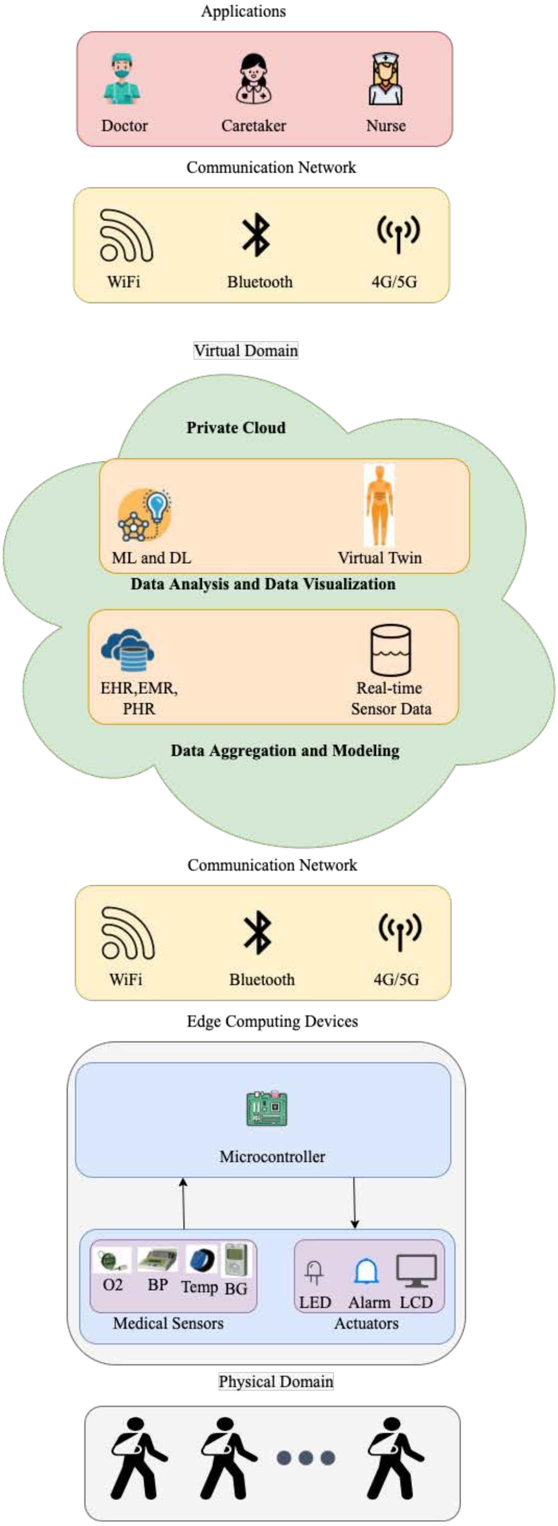


Figure 6-3: Digital Twin Architecture for In-Patient Monitoring System.

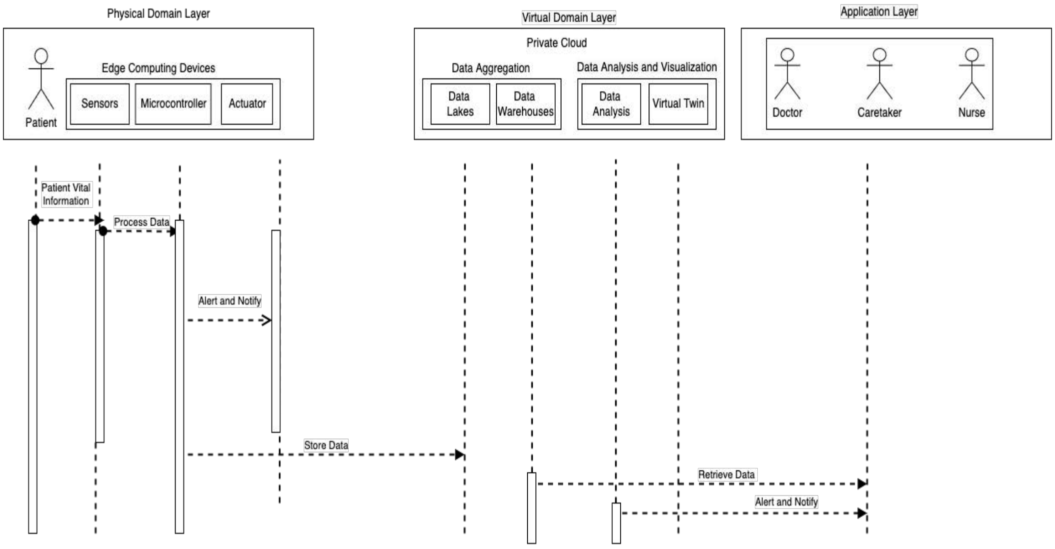


Figure 6-4: Sequence Diagram for In-Patient Monitoring System.

6.3.1.2 Sniffing

It is a passive attack in which the attacker tries to learn the communication that takes place between two entities. One way to launch this attack is by installing a software called Wireshark that helps to sniff the packets. In addition to this, it is also important for the attacking machine to use Wireshark on the same local network as the microcontroller. After confirming this, the attacker will first perform an analysis of security on remote access and on the network to which the microcontroller is connected. Then through Wireshark, the attacker will sniff the packets on the network that the microcontroller is receiving. Figure 6-5 illustrates a visual representation of this attack.

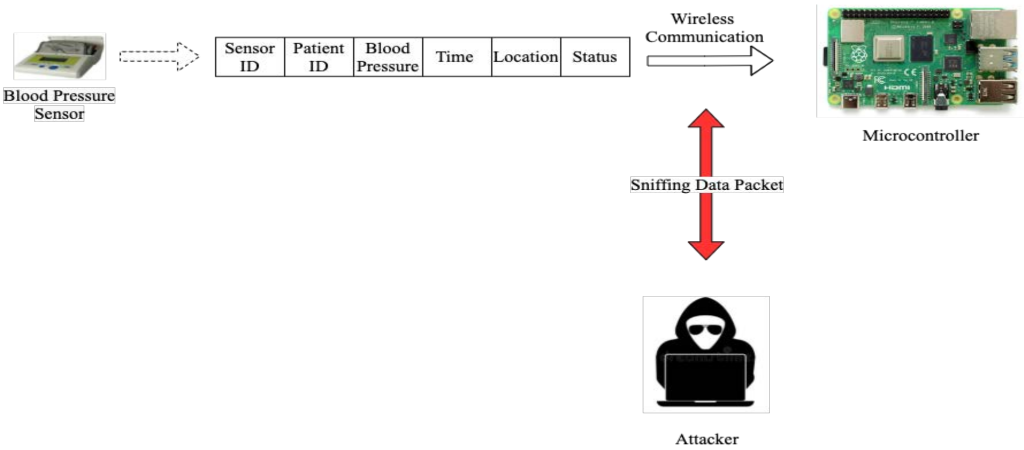


Figure 6-5: Sniffing Attack.

In terms of the in-patient monitoring system, the attacker will use this attack to sniff the sensitive information of the patient like their vital information that were collected via the sensors.

6.3.1.3 Denial of Service (DoS)

It is a network-based attack whose main purpose is to disable the access to the target network. This attack can only be implemented on the devices that are connected on the same local network. To launch this attack, the attacker first manipulates the received data packets on the network using Scapy within Python. Scapy is a packet manipulation tool that lets users to send altered packets over the network. This software can therefore be used to flood the network with large amounts of packets. This will thus leave the microcontroller unable to handle the massive data packets as received from the sensors. Figure 6-6 illustrates an overview of the DoS attack. This attack when employed in the in-patient monitoring system will disable the access to the communication that exists between the sensor and the microcontroller.

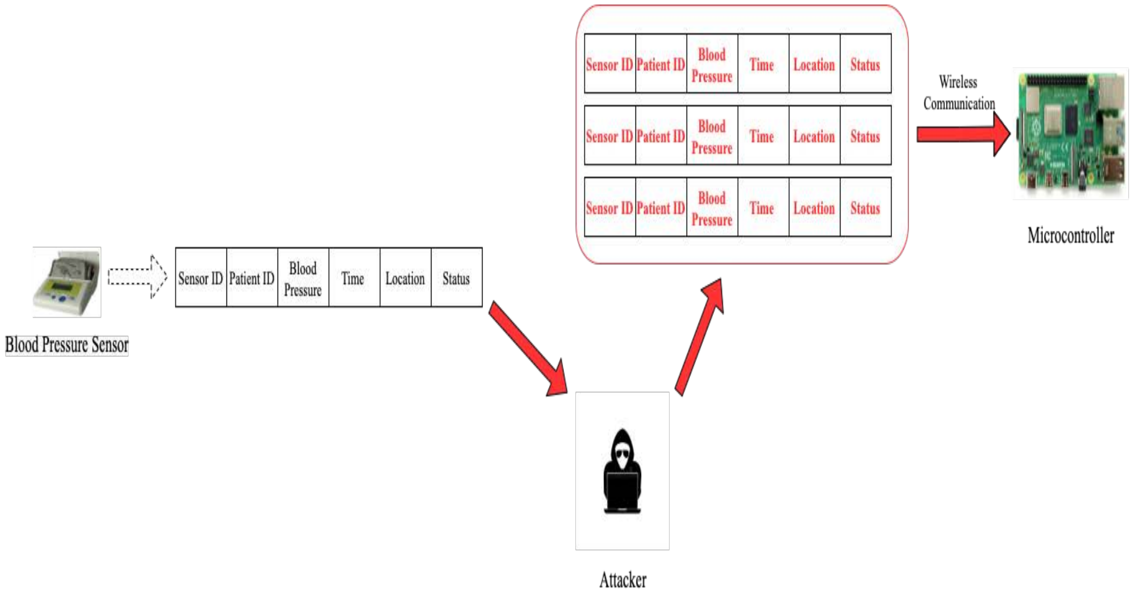


Figure 6-6: DoS Attack.

6.3.2 Physical domain layer to virtual domain layer

The data packets received from the sensors gets collected and converted to a single data packet at the DT. However, once this data is received at the DT, it needs to be forwarded further to the cloud to facilitate the required operations. That is, as illustrated in Figure 6-7, data now flows from the data acquisition unit to the cloud. The role of the cloud in this process is to initially store the collected data in the respective data repositories (data lakes and data warehouses). However, just like the previous data flow process, this data flow process can also be vulnerable to a variety of security attacks. To narrow down the scope of this thesis, generated Virtual Twin is considered to be secure enough to mitigate any security attacks faced. Thus, mentioned below are some of the major types of security attacks that can be faced while storing the patient information in the cloud:

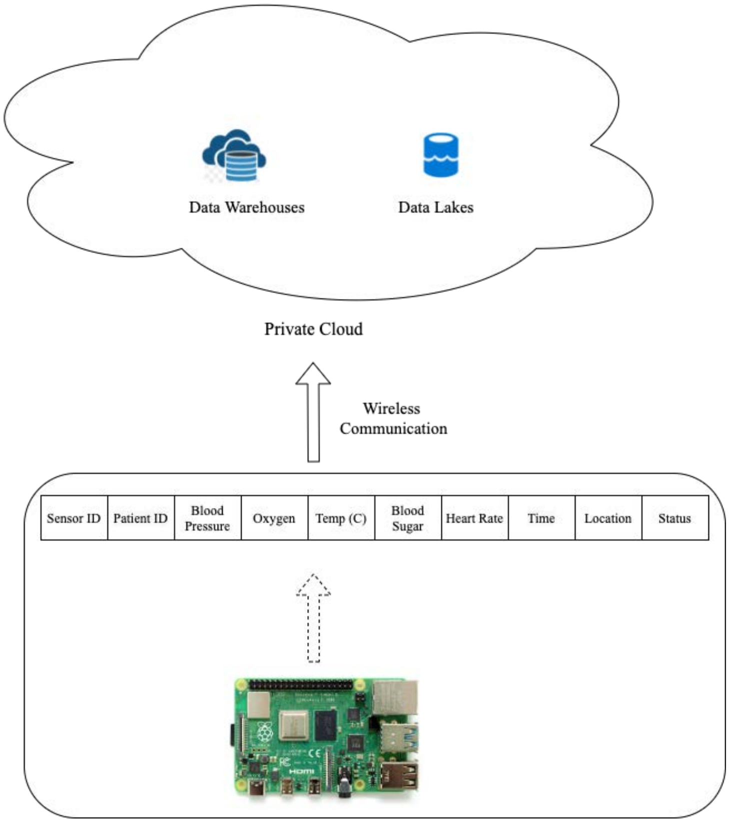


Figure 6-7: Data Flow from Physical Domain Layer to Virtual Domain Layer.

6.3.2.1 Data modification

This attack can be launched by anyone who has the access to the data repositories present in the cloud where the sensitive information might be stored. The attacker in this attack will manipulate the data and extract confidential information from the inside as well. Figure 6-8 depicts a visual representation of this attack in the context of the in-patient monitoring system. As seen from the figure, the attacker tries to manipulate any of the vital information of the patient like blood pressure, blood sugar, oxygen, temperature, or they could also change the time, health status, patient ID, sensor ID, and location of the patient.

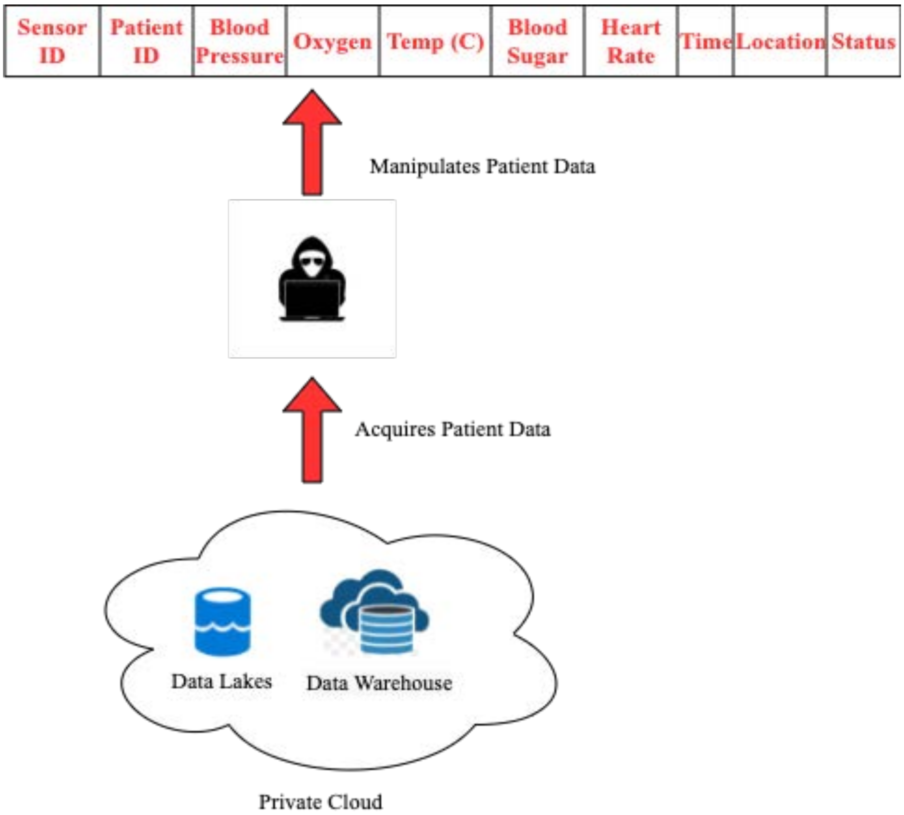


Figure 6-8: Data Modification Attack.

6.3.2.2 SQL injection

This type of attack can be launched only on the data stored in the data warehouse as this attack requires the data to be structured to execute SQL commands. Therefore, in order for the attacker to launch such the SQL injection attack, the attacker needs to inject SQL commands in a program that would allow the attacker to acquire the confidential information or data of any user and even manipulate the records stored in the database if needed. If the given database is not secure, then it will be able to execute these malicious commands that would help to facilitate in providing the attacker access to the stored information or data. In case of the in-patient monitoring system, the attacker can launch this attack to gain the sensitive information of the patient. Figure 6-9 illustrates a visual representation of the SQL injection attack. As seen from the below illustrated figure, the attacker can gain access to the vital information of the patient such as their ID, blood pressure, blood sugar, oxygen, temperature, health status, location if it is able to execute the SQL injection attack successfully.

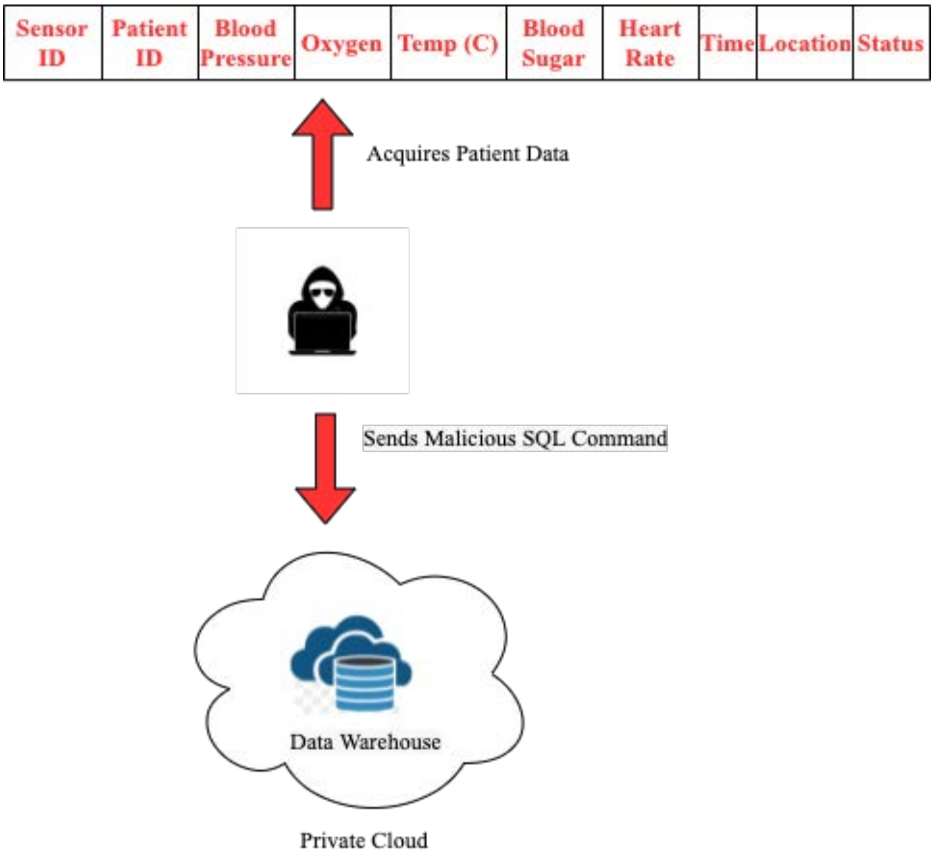


Figure 6-9: SQL Injection Attack.

6.3.2.3 Unauthorized attack

It is a type of attack, that allows an unauthorized attacker to infiltrate the system and prevent the access of the actual authenticated users into the system. In addition to this, the attacker can also delete any of the sensitive information completely if they want to do so. Figure 6-10 illustrates a visual representation of this attack. As seen from the figure, the attacker can gain access in order to manipulate the patient data. Whereas Figure 6-11 illustrates how an attacker through the execution of this attack can prevent the healthcare user (doctor, nurse, caretaker) to access the respective patient data by deleting the stored patient data.

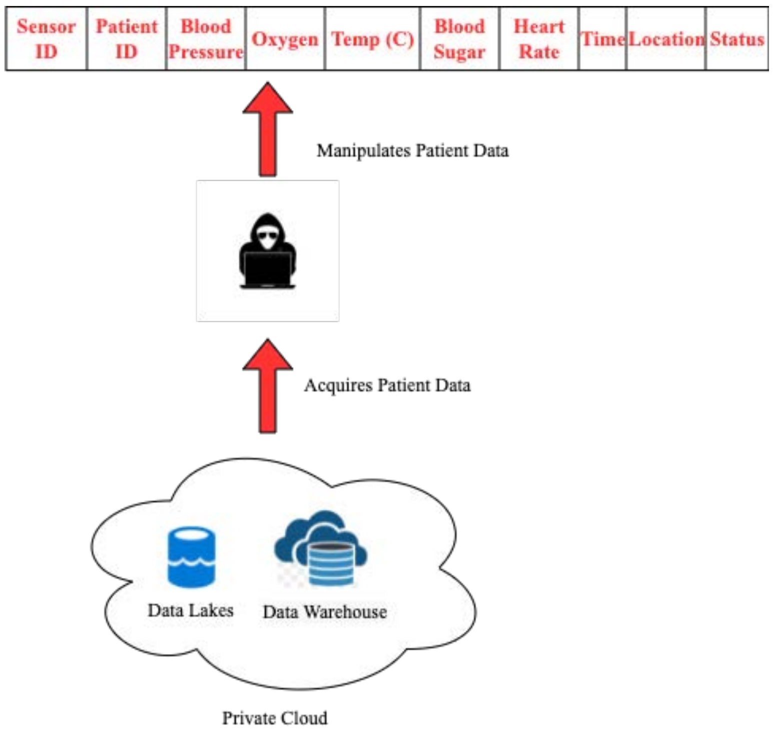


Figure 6-10: Unauthorized Attack where Attacker Prevents Patient Data Access.

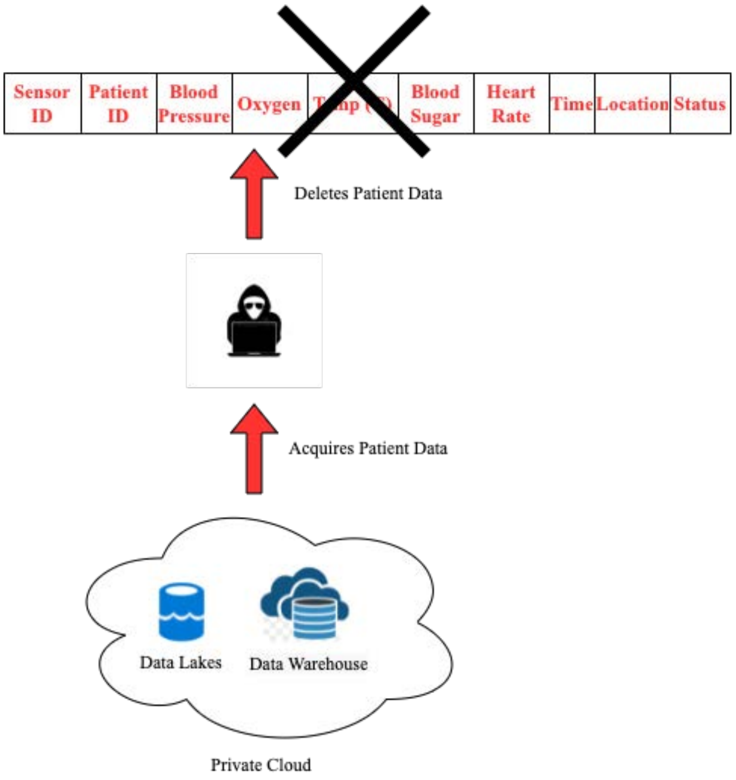


Figure 6-11: Unauthorized Attack where Attacker Deletes Patient Data.

6.3.3 Virtual domain layer to application layer

After the data of the patient is stored in one of the data repositories located in the cloud, the next step is to convert this data into some useful form. That is, the received data is first analysed and then visualized. During the visualization process, a respective Virtual Twin is generated. This Virtual Twin depends on the application it is being used for. Figure 6-12 depicts the data flow process that takes place between the virtual domain layer and application layer. In the context of the healthcare domain, it could be anything ranging from a device to the entire body of the respective patient. Once this Virtual Twin is generated, it is the accessed by the respective users. In this case, it could be a doctor, nurse, or a caretaker. The doctor can use the Virtual Twin to determine efficient patient diagnosis. While the nurse and caretaker can use the Virtual Twin to monitor the current health status of their patients. However, the data when being accessed by the users can be vulnerable to a variety of security attacks. Thus, mentioned below are some of the major types of security attacks that can be faced by the users:

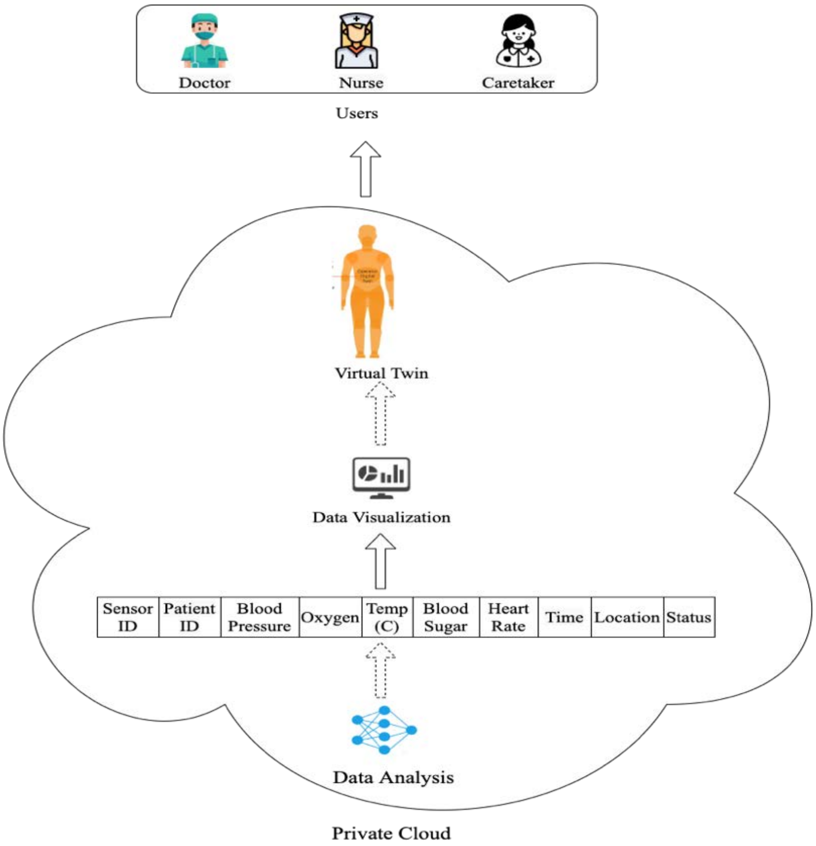


Figure 6-12: Data Flow from Virtual Domain Layer to Application Layer.

6.3.3.1 Session hijacking

The attacker exploits the authentication protocols and alters the session management of the network. In this way, the attacker gains access to the personal information of the user and use the network just like the real user. Figure 6-13 illustrates a visual representation of this attack.

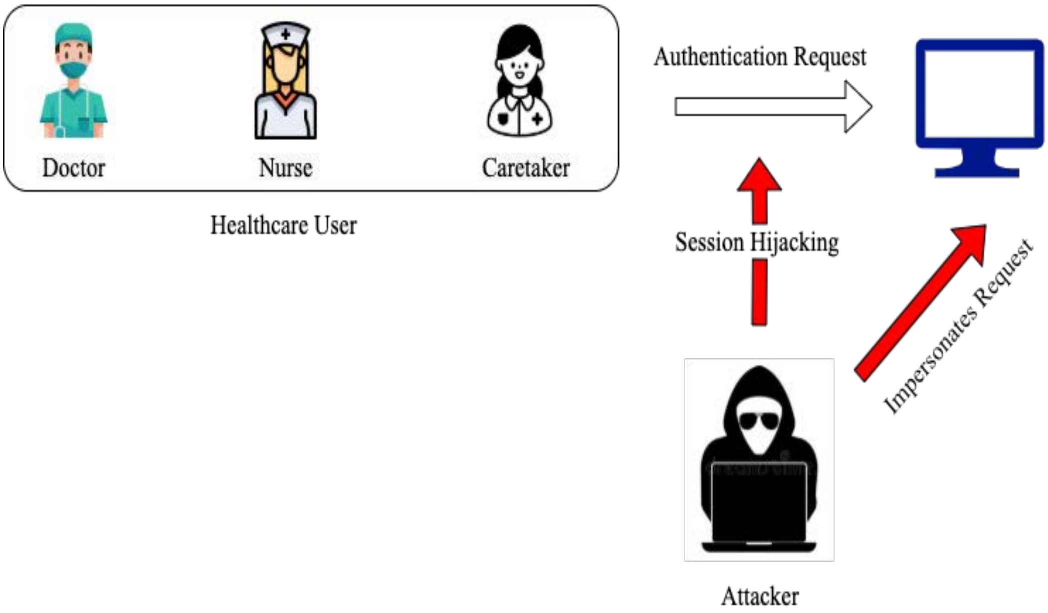


Figure 6-13: Session Hijacking.

6.3.3.2 Cross-site Scripting (XSS)

This attack works on compromising the trust relationship between the user and the web application site by injecting malicious code. Attackers can control the integrity of the application. Hence, the security and the privacy of the disabled users will be breached. Figure 6-14 illustrates a visual representation of this attack.

6.3.3.1 Phishing

In this type of attack, the attacker pretends to be a legitimate user or institution to obtain sensitive information about the users, such as passwords and credit card details. The common medium for this attack is email, where sensitive information is acquired by an attacker when the users open the illegitimate email. Figure 6-15 illustrates a visual representation of this attack.

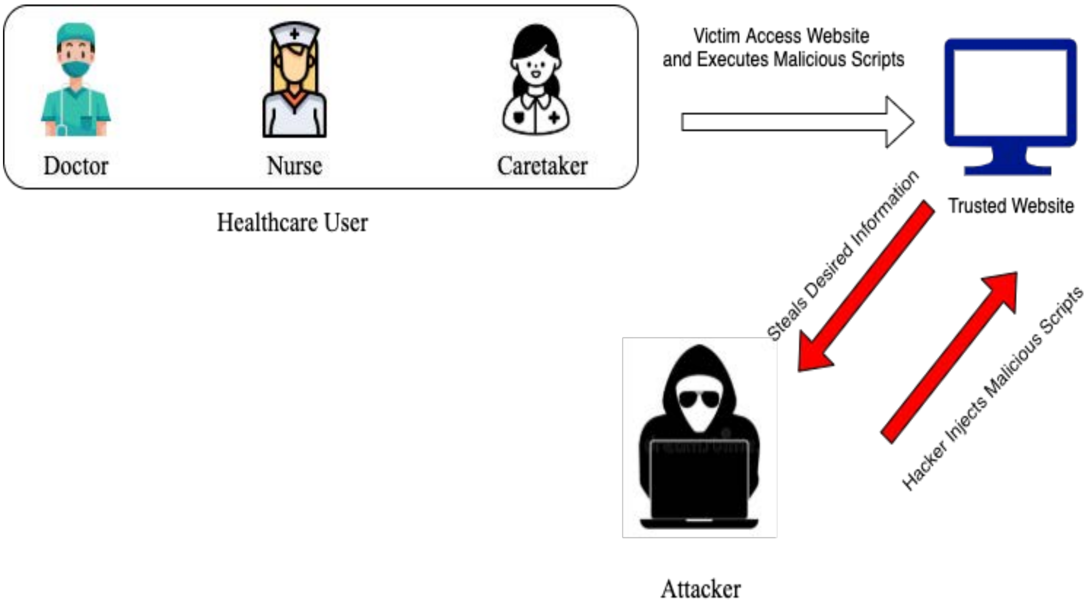

Doctor
Nurse
Caretaker
Healthcare User
Victim Access Website and Executes Malicious Scripts
Trusted Website
Steals Desired Information
Hacker Injects Malicious Scripts
Attacker

Figure 6-14: Cross-Site Scripting (XSS) Attack.

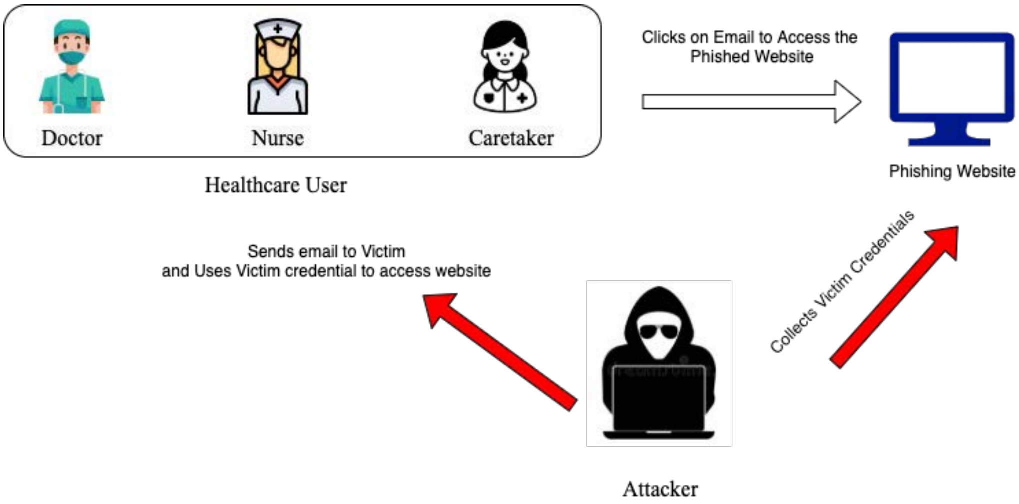

Doctor
Nurse
Caretaker
Healthcare User
Clicks on Email to Access the Phished Website
Phishing Website
Sends email to Victim and Uses Victim credential to access website
Collects Victim Credentials
Attacker

Figure 6-15: Phishing Attack.

Chapter 7. Experimental Work

In this chapter, experimental scenarios that were used to perform the implementation of the proposed work were presented.

7.1. Formal Modelling of DT System

A simple case study for in-patient monitoring system in DT as illustrated in Figure 7-1 is used as an example to implement the proposed DT security architecture. For this case study, a smart hospital is considered that can compromise of multiple in-patients which are connected to various medical sensors to measure vital signs like blood pressure, blood glucose, heart rate, temperature, oxygen level, etc. The sensors are connected to the patient through the wireless link labelled as L1. Once the data is collected it is processed by the microcontroller that is connected by another wireless link L2. Then a gateway is used to connect the microcontroller to the cloud where the data is aggregated, modelled, and analysed to obtain the corresponding Virtual Twin. After the Virtual Twin is obtained, it is then accessed by different healthcare users like doctor, nurse, caretaker using an application, web interface that is connected via a wireless connection through the link labelled as L5. In addition to this, the doctor, nurse, or caretaker can also receive any changes related to the patient health status. Table 7-1 illustrates a tabular form of the different notations that were used in this equation.

Table 7-1: Summary of the Notations Used.

DT Architecture Layer	Components	Notation Used
Physical Domain Layer (Requires Vicinity Access)	Microcontroller, sensors, link 1, link 3 with microcontroller level encryption	$MC, SEN, L1, L3$
Communication Network Layer (Requires Vicinity Access)	Gateway	GW
Virtual Domain Layer	Data centre	DC
Application Layer	Application, link 5 with application level encryption	$APP,\ L5$

In terms of the security mechanisms present in each individual of these mentioned component, links L1 to L5 is considered to compromise of a very weak encryption mechanism like WEP [99] [100]. On the other hand, the edge computing devices, that is, the microcontroller, sensor, and gateway comprises of a medium-level cryptographic algorithm like RC4 algorithm [101]. However, for the case of the data centre, a very strong firewall is considered for its protection [102], and finally the application is protected by a very weak web application firewall like citrix [103]. Illustrated below shows the increasing order of the security mechanisms present for each component:

$$DC_{security} > MC_{security}, GW_{security}, SEN_{security} > APP_{security} > LINKS_{securtiy}$$

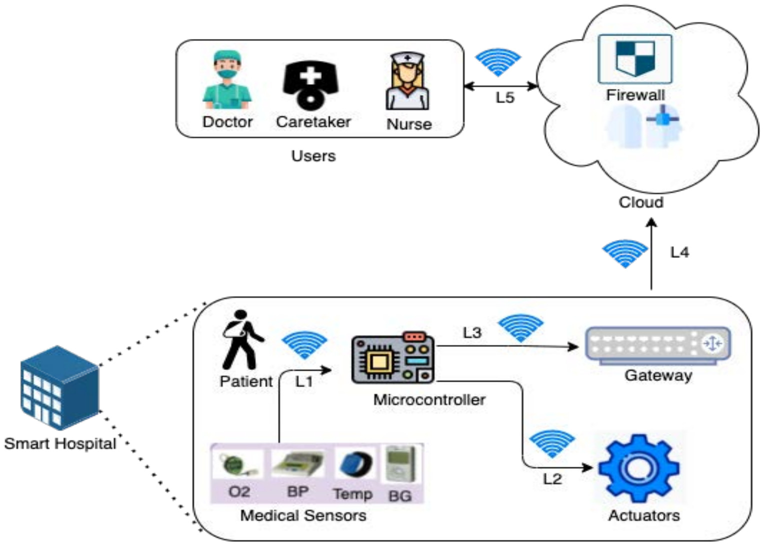


Figure 7-1: In-Patient Monitoring System in DT.

However, just like any typical case study, this case study comprises of different components and wireless links that can be vulnerable to various security attacks. Therefore, four common security attack objectives were chosen based on how often they can occur through the extensive analysis of research papers conducted. Table 7-2 illustrates some of the different research papers that mention about these attacks. Even though all of these attacks are commonly found in many research papers, the ransomware attack amongst them is one of the most common attacks as it has become increasingly popular over the past few years. Furthermore, Table 7-3 shows how these attacks can compromise the different security violations. Thus, the upcoming subsections, talks in brief detail regarding the following chosen attacks:

Table 7-2: Common Attack Objectives.

Attack Objective	References
Data Access	[104], [105], [106], [107], [108]
Data Modification	[108], [109], [110], [111], [112]
DoS	[113], [114], [115], [116], [117]
Ransomware	[118], [119], [120], [121], [122]

Table 7-3: Summary of Attack Objectives.

Attack Objective	Security Violations
Data Access	Confidentiality
Data Modification	Integrity, Availability
DoS	Availability
Ransomware	Integrity, Availability

7.1.1 Data Access (DA)

To achieve this attack objective, the attacker would launch some form of a passive attack which helps in accessing the patient data. This attack can be targeted towards one or more components present in the respective DT architecture as demonstrated in Equation-12:

$$DA = \begin{cases} C_{DC} & data\ center\ compromised \\ C_{APP} & application\ compromised \\ C_{L5} & given\ encryption\ (app) = \perp \\ C_{GW} \wedge VA & gateway\ comrpromised \\ C_{L1} \wedge VA & Link\ L\ 1\ compromised \\ C_{MC} \wedge VA & microcontroller\ compromised \\ C_{L3} \wedge VA & given\ encryption\ (MC) = \perp \end{cases} \tag{11}$$

From Equation-12, it is apparent that the attacker can achieve its objective through multiple ways. For instance, it can target the components present in the communication network layer like the gateway (GW) or it can also target components present in the physical domain layer like the ($L1$) link or the micro controller (MC) or the link ($L3$) that comprises of a micro controller level encryption (MCE). However, it is important to note that the launch of any attacks on the components present in both of these layers can only be possible if the attacker has vicinity access (VA). This access can be gained by the attacker through various reasons like impersonating oneself as an authorized person, being an insider attacker, knowing someone who has access to the system in order to gain access, etc. In terms of targeting the components present in the virtual domain layer, the attacker can compromise the data centre (DC) which stores the respective patient data. Finally, in the application layer the attacker would need to compromise the respective link ($L5$) and the application-level encryption present in this layer (AE) to gain access to the data or it can just launch an attack directly on the user application (app) itself. In terms of complexity for this attack objective, since this attack is a passive form of attack, therefore, the complexity of such an attack would be considered as low [123].

7.1.2 Data Modification (DM)

As opposed to the earlier attack objective, in this case the attacker wants to modify the patient data. Just like the earlier objective, the attacker in this case can also launch its attack by targeting one or more components present in the DT architecture. Therefore, Equation-13 illustrates the different ways an attacker can compromise one or more components present in the DT architecture:

$$DM = \begin{cases} C_{DC} & data\ center\ compromised \\ C_{APP} & application\ compromised \\ C_{SEN} \wedge VA & sensor\ compromised \\ C_{MC} \wedge VA & microcontroller\ compromised \\ C_{GW} \wedge VA & gateway\ compromised \end{cases} \tag{12}$$

For instance, the attacker can try to compromise the data centre (DC) in order to modify the patient data stored in it or it can try to compromise the application (app) so that it can modify the patient data that the application uses to display to the healthcare user or in terms of the physical layer components such as the gateway (GW) or the micro controller (MC) or the sensor (sen) as all of these components also collect and store patient data to send it to the upcoming layers to facilitate the further processing of data. However, it is important to note that for the accessing the components in the physical domain layer, the attacker requires access to the vicinity (VA). For the complexity of this attack, this attack is considered as a high level of complexity attack. This is because in this case the attacker first needs to gain access and also ensure the modification of the data, which requires the acquisition of higher privilege level by the attacker in order to be able to modify the current data [124].

7.1.3 Denial-of-Service (DoS)

To achieve this attack objective, the attacker would try to flood the targeted components with large amount of traffic or information to make it inaccessible for its intended use. For this case, to achieve this objective, the attacker can try to target any of the components or links present in the DT system. Thus, the attacker can target the links $L1, L2, L3, L5$ or the components $GW, MC,$ or DC as demonstrated in Equation-14:

$$DoS = \begin{cases} C_{GW} \wedge VA & gateway\ compromised \\ C_{L1} \wedge VA & link\ 1\ compromised \\ C_{MC} \wedge VA & microcontroller\ compromised \\ C_{L2} \wedge VA & link\ 2\ compromised \\ C_{L3} \wedge VA & link\ 3\ compromised \\ C_{DC} & data\ center\ compromised \\ C_{L5} & link\ 5\ compromised \end{cases} \tag{13}$$

In terms of complexity for this attack, the execution of this attack would be considered as a medium level security level, as in this case the attacker has to not only gain access to the targeted components, but it also needs to send a large amount of information to these components [125].

7.1.4 Ransomware (RW)

To achieve this attack objective, the attacker would try to insert a form of malicious software to the targeted component in order to lock and encrypt the confidential information. This information can only be returned back once the demanded ransom is paid to the attacker. Therefore, in such an attack, the attacker would target the critical components present in the given system. In this case, it would be the data centre (DC). The data centre would be compromised by following numerous steps that needs to be executed in sequential order as shown in Equation-15:

$$RW_{DC} = \begin{cases} Access \\ Super - user\ Access \\ Exfiltration \\ Encryption \end{cases} \tag{14}$$

From Equation-15 it is apparent that the ransomware attack cannot be executed in a single step unlike the execution of the previous attacks. A summary for the execution of these attacks has been described as below [126]:

1. Component Access: The first step for the execution of the attack is that just like in any attack, the attacker requires access to the targeted component.
2. Privilege Escalation to Super-User: After the attacker successfully gains access of the targeted component, the next step is to elevate its permission in order to perform lateral movements as well as to gain access to the stored data.
3. Exfiltration: Once the attacker gains super user privilege, the next step is to steal the stored data and to transfer the data to itself.
4. Encryption: Finally, the collected data is locked using an encryption mechanism after which the data is not available to the authorized user. A ransom is then demanded to unlock the data and gain access to it for the authorized user.

Furthermore, to generate the results, the parameter values depicted in Table 7-4 were used as input to the proposed model. Based on the given in-patient monitoring system scenario, the attacker can carry out a variety of attacks. The attacker in this case can be classified mainly into a naive attacker and a skilful attacker. To explore a variety of possibilities for the chosen healthcare case study, the obtained results were organized according to the following three major categories:

1. Vicinity Access: In this case, the attacker is considered to have access to all the components that require vicinity access to execute the chosen attack.
2. No Vicinity Access: For this case, the attacker is considered to only have access to all the components that do not require vicinity access in order to execute the chosen attack.
3. Both: Here, the attacker has access to both the components that require vicinity access and that does not require vicinity access to launch an attack.

Table 7-4: Parameter Values Used in Prism Models.

Parameter	Values
Attack Objective	DA, DM, DoS, RW
Probability of Attack Success (DA, DM, DoS, RW)	0.9, 0.3, 0.5, 0.2

Probability of Defence Profile Levels (High, Medium, Low)	0.594, 0.51, 0.432
Vicinity Access (VA)	0-1
Microcontroller-Level Encryption	True (for β_h and β_m)
Application-Level Encryption	True (for β_h and β_m)

Chapter 8. Results Discussion

In this chapter, the various results obtained through the implementation of the experimental scenario mentioned from the previous chapter is demonstrated.

8.1. Results

After using the parameter values mentioned in the previous chapter as input to the proposed model, the following results were obtained for the below attack objectives:

8.1.1 Data Access (DA)

In case of the DA attack objective, the naïve attacker has 3 options, that is, it has the option to choose the components that require vicinity access component, or to choose the components that do not require vicinity access or to choose all the components. The results for these three cases were generated as demonstrated in Figure 8-1. On the other hand, Table 8-1 demonstrates a tabular representation of the values obtained. Both Figure 8-1 and Table 8-1 demonstrates the result for the case when the system comprises of high -level defence. Whereas Figure 8-2, Figure 8-3 and Table 8-2 and Table 8-3 represents the medium-level and low-level defence for this case. Therefore, through the references of these tables and figures, the following observations were found:

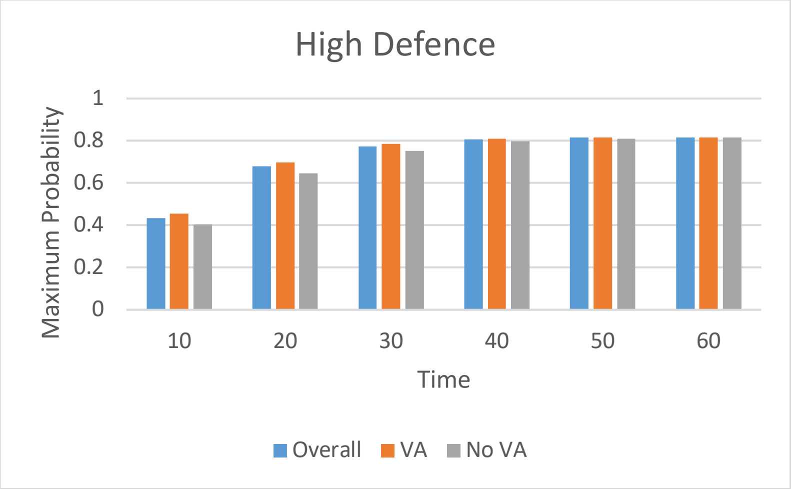


Figure 8-1: Probability of DA Success for High Defence.

Table 8-1: Probability of DA Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.435	0.454	0.404
20	0.679	0.698	0.646
30	0.774	0.785	0.753
40	0.806	0.810	0.796
50	0.814	0.815	0.811
60	0.816	0.816	0.815

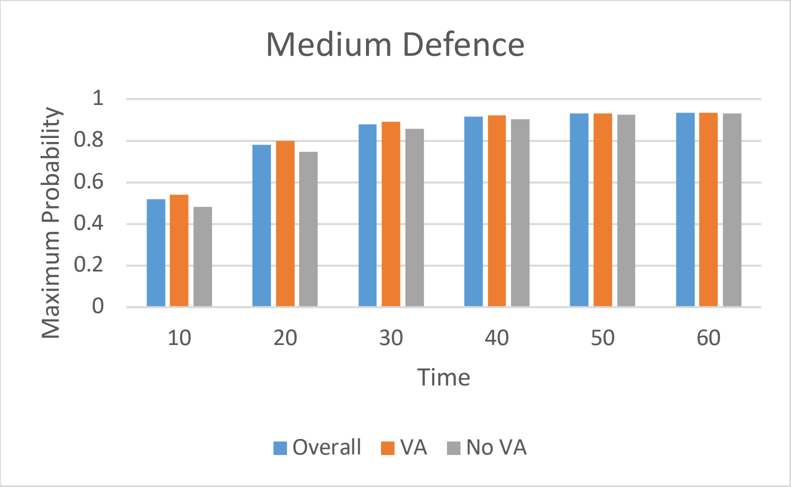


Figure 8-2: Probability of DA Success for Medium Defence.

Table 8-2: Probability of DA Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.517	0.539	0.481
20	0.779	0.798	0.745
30	0.878	0.889	0.855
40	0.916	0.922	0.903
50	0.929	0.931	0.923
60	0.932	0.933	0.930

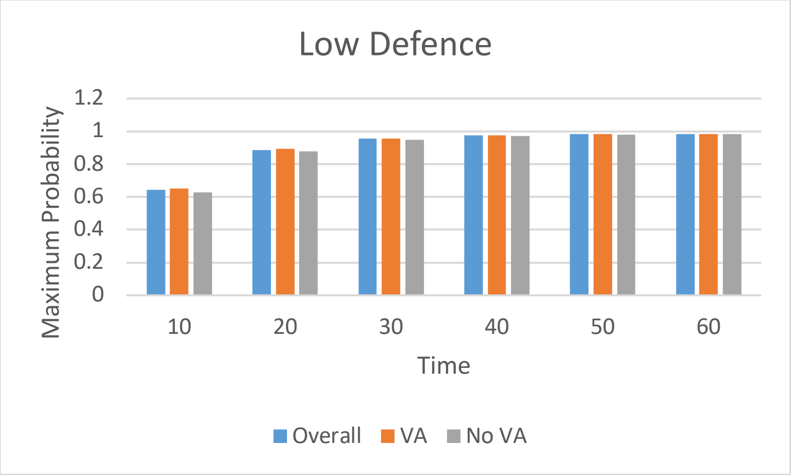


Figure 8-3: Probability of DA Success for Low Defence.

Table 8-3: Probability of DA Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.642	0.651	0.629
20	0.887	0.894	0.878
30	0.954	0.957	0.950
40	0.975	0.976	0.973
50	0.981	0.982	0.981
60	0.983	0.983	0.983

1. The highest probability of objective success is for the low defence. This is then followed by the probability of success for the medium defence and the lowest probability of success is for the high defence as it is the toughest to launch an attack for this case.
2. For the case when the attacker has VA, the probability of attack success is higher than for the case when the attacker has no VA, this is because in this case the attacker has more chances to attack the weakest security component. That is, for this case, the attacker can attack the link $L1$ and $L3$, as opposed to only having the choice to attack on a single weak component ($L5$).

The results obtained for the attack objective cost in terms of the three defence-level can be inferred from Figure 8-4 with Table 8-4, Figure 8-5 with Table 8-5, Figure 8-6 with Table 8-6. Through the references of all these tables and figures, the following observations were made:

1. In the case when the attacker has no VA, as shown in Figure 8-6 and Table 8-6, although the cost for the high and medium defence has some difference amongst them, however, in the case of the low defence a significant lesser difference is required by the attacker. This is because in this case the attacker does not have to deal with encryption for the link $L5$ which in turn requires lesser effort in case these components are chosen.
2. For the case when the attacker has VA, although the attacker has more chances to attack the least secure component ($L1$), however in this case the attacker requires more cost for gaining VA which nullifies this advantage and thereby requires more cost.
3. In the case, where the attacker has access to all the components, the attack cost is slightly more as compared to the cost when the attacker has VA. This is because in this case, although the attacker has access to all the available links but in addition to this it also the option to launch an attack on the DC which requires more effort as compared to when the attacker only has VA.

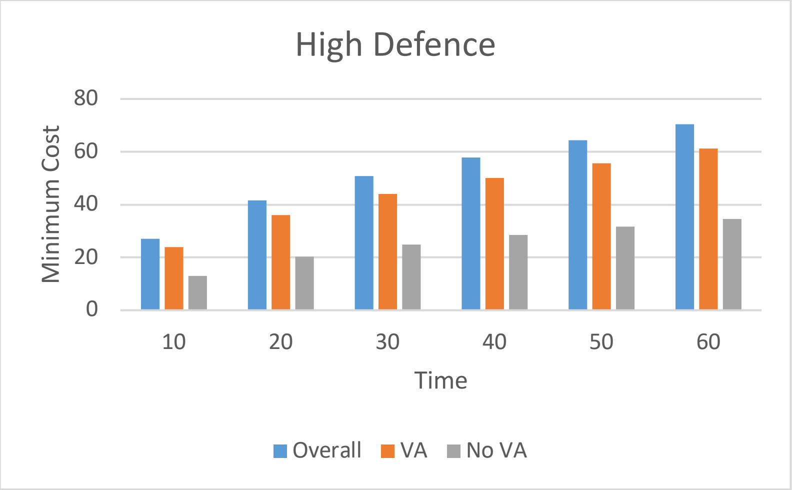


Figure 8-4: Minimum DA Cost for High Defence.

Table 8-4: Minimum DA Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	27.141	23.849	12.930
20	41.494	36.026	20.197
30	50.809	43.916	24.970
40	57.911	50.010	28.541
50	64.294	55.616	31.634
60	70.515	61.150	34.574

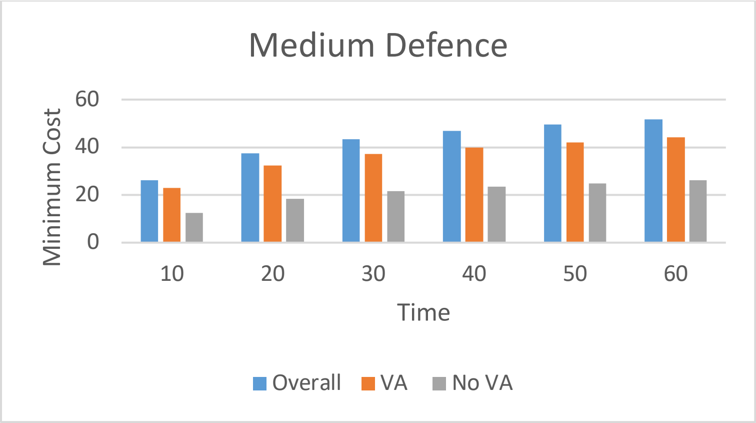


Figure 8-5: Minimum DA Cost for Medium Defence.

Table 8-5: Minimum DA Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	26.082	22.866	12.470
20	37.432	32.295	18.398
30	43.306	37.071	21.602
40	46.828	39.935	23.532
50	49.439	42.137	24.903
60	51.759	44.166	26.049

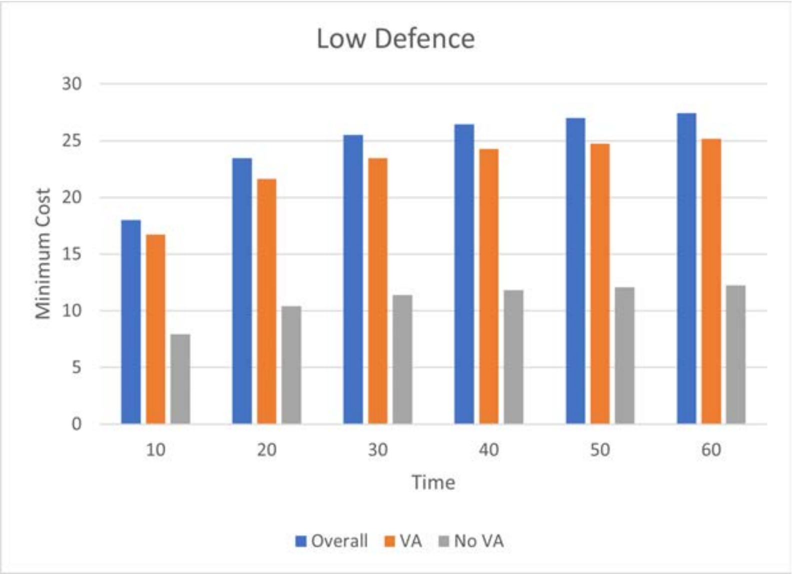


Figure 8-6: Minimum DA Cost for Low Defence.

Table 8-6: Minimum DA Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	18.024	16.727	7.921
20	23.485	21.651	10.422
30	25.533	23.466	11.381
40	26.443	24.268	11.812
50	26.990	24.759	12.064
60	27.431	25.165	12.261

For the skilful attacker who chooses the DA as its objective, the results obtained for the probability of success in terms of the three defence-level can be inferred from the result generated in Figure 8-7 with Table 8-7, Figure 8-8 with Table 8-8, Figure 8-9 with Table 8-9. Therefore, through the reference of these figures and tables the following observations were made:

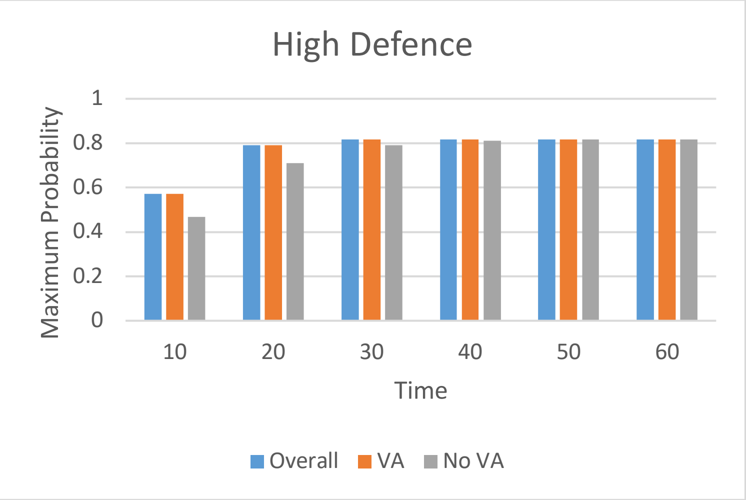


Figure 8-7: Probability of DA Success for High Defence.

Table 8-7: Probability of DA Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.572	0.572	0.467
20	0.790	0.790	0.710
30	0.816	0.816	0.791
40	0.816	0.816	0.812
50	0.816	0.816	0.816
60	0.816	0.816	0.816

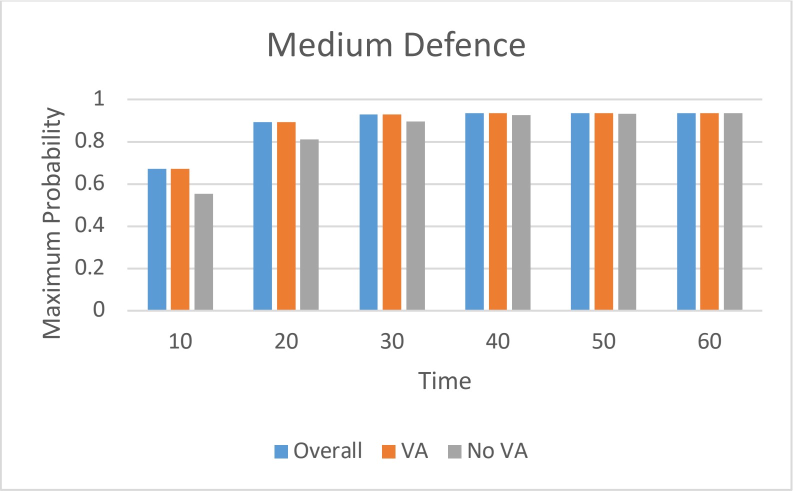


Figure 8-8: Probability of DA Success for Medium Defence.

Table 8-8: Probability of DA Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.670	0.670	0.554
20	0.892	0.892	0.811
30	0.929	0.929	0.896
40	0.933	0.933	0.925
50	0.933	0.933	0.932
60	0.933	0.933	0.933

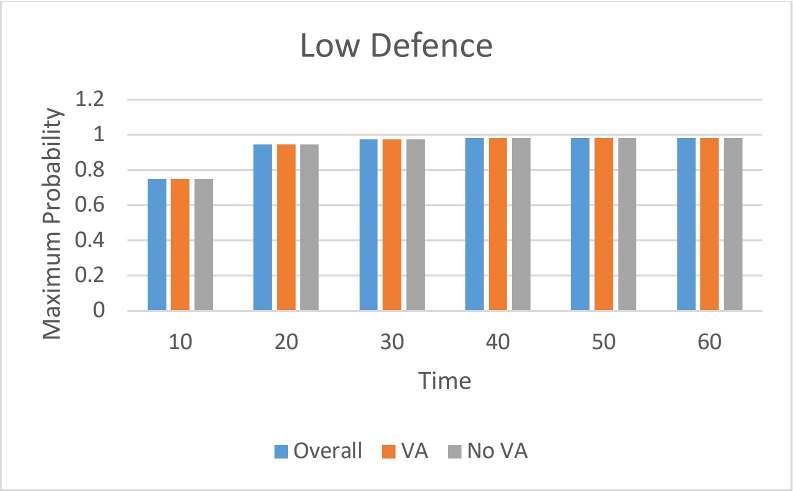


Figure 8-9: Probability of DA Success for Low Defence.

Table 8-9: Probability of DA Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.751	0.751	0.751
20	0.947	0.947	0.947

30	0.975	0.975	0.975
40	0.983	0.983	0.983
50	0.983	0.983	0.983
60	0.983	0.983	0.983

1. For the case when the attacker has no VA, the attacker performs comparatively worse in terms of probability of success as shown in Figure 7-8. The reason for this case is that here the attacker has only the option of link $L5$ that has the least security as compared to the other components available. And since $L5$ is secured with encryption, it is difficult to execute the data access attack in this as compared to $L1$. However, for the case when the defence is low, the probability of success is the same for all the cases. This is because in this case, $L5$ is not encrypted, and this thereby makes it equal to attacking it on $L1$.
2. Since in this case the attacker is advanced in terms of skills and experience, therefore, the attacker would therefore make planned decisions on which components to launch its attack and thus gain the maximum probability of success with the least cost possible. Thus, for both cases of VA and where attacker has access to all the components, the attacker attains the same probability of success as well as the cost of attack. The reason for this is because since the attacker is a skilful one, the attacker is therefore smart enough to choose the link $L1$ that has the least security amongst the other available components. Therefore, for both VA and when the attacker has access to all components, the attacker chooses the $L1$ component and thereby yielding the same results for probability of success and cost.

The results obtained for the cost in terms of the three defence-level can be inferred from Figure 8-10 with Table 8-10, Figure 8-11 with Table 8-11, Figure 8-12 with Table 8-12. Through the references of all these tables and figures, the following observations were made:

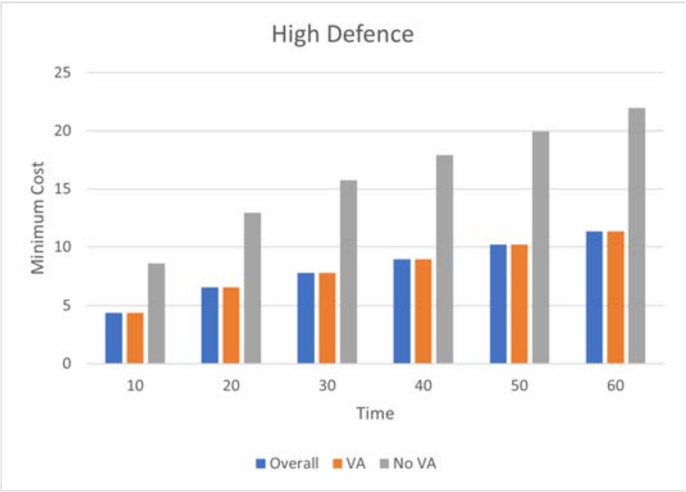


Figure 8-10: DA Cost for High Defence.

Table 8-10: DA Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	4.378	4.378	8.632
20	6.577	6.577	12.945
30	7.817	7.817	15.742
40	8.987	8.987	17.919
50	10.204	10.204	19.957
60	11.333	11.333	21.982

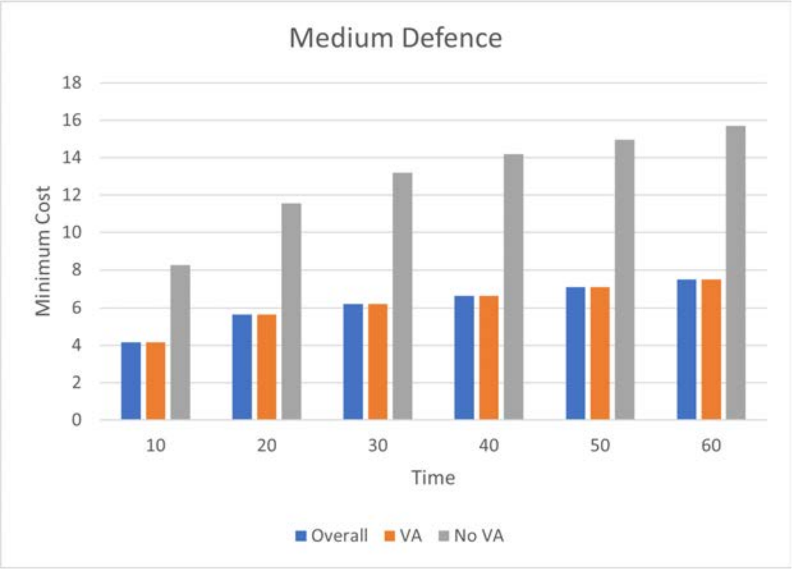


Figure 8-11: DA Cost for Medium Defence.

Table 8-11: DA Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	4.151	4.151	8.263
20	5.635	5.635	11.552
30	6.215	6.215	13.197
40	6.644	6.644	14.185
50	7.082	7.082	14.967
60	7.488	7.488	15.704

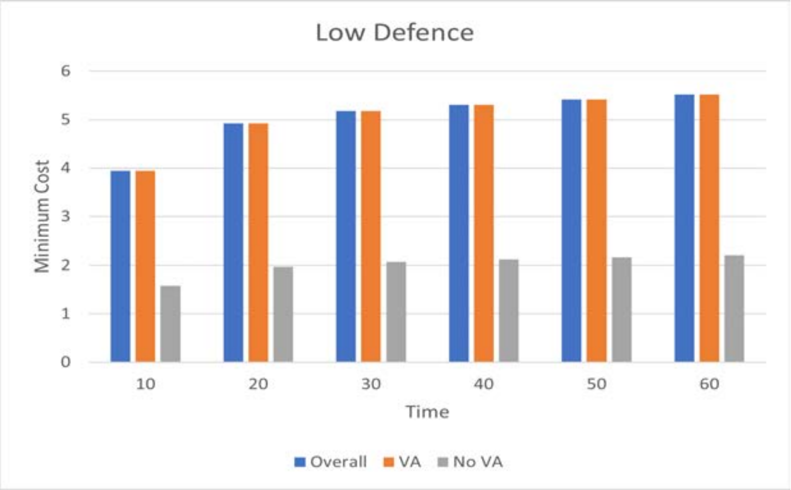


Figure 8-12: DA Cost for Low Defence.

Table 8-12: DA Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	3.941	3.941	1.576
20	4.922	4.922	1.968
30	5.182	5.182	2.072
40	5.309	5.309	2.123
50	5.417	5.417	2.167
60	5.517	5.517	2.207

1. For the cost required in the case when the attacker has no VA, it has been found that the attacker requires more cost as compared to the naïve attacker cost for high and medium level defence. This is because of the encryption mechanism present in $L5$. However, the cost for low defence is much lesser than the naïve attacker case, this is because besides $L5$ and $L1$ having the no encryption, in addition to this, the link $L5$ does not require VA which reduces the effort by the attack even more as compared to $L1$ that requires VA.

2. The cost is the same for the case when the attacker has access to all the components or only the VA components, the reason for the similar cost for both the case is because the attacker in this case is a skilful one, therefore, the attacker is smart enough to choose the $L1$ that has the least security amongst the other available components.

8.1.2 Data Modification (DM)

In the case of DM for a naïve attacker, the results obtained for the probability of success in terms of the three defence-level can be inferred from Figure 8-13 with Table 8-13, Figure 8-14 with Table 8-14, Figure 8-15 with Table 8-15. Through the references of all these tables and figures, the following observations were made:

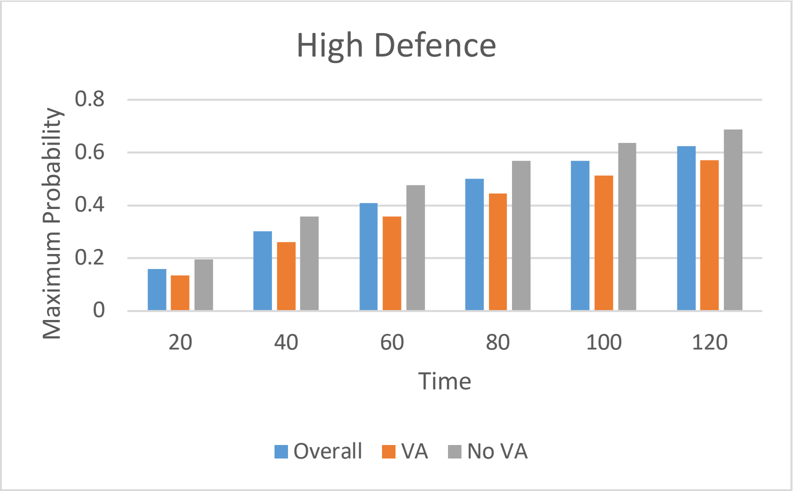


Figure 8-13: Probability of DM Success for High Defence.

Table 8-13: Probability of DM Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.160	0.135	0.195
40	0.302	0.260	0.358
60	0.409	0.358	0.476
80	0.500	0.445	0.569
100	0.569	0.513	0.636
120	0.625	0.570	0.686

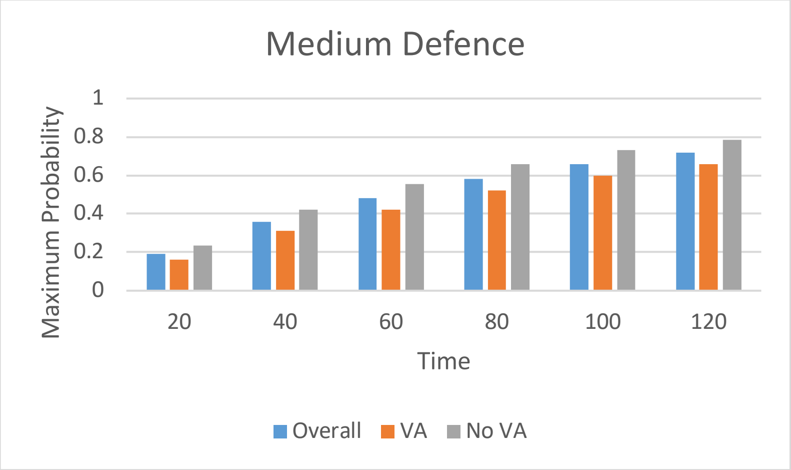


Figure 8-14: Probability of DM Success for Medium Defence.

Table 8-14: Probability of DM Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.191	0.162	0.233
40	0.358	0.310	0.422
60	0.480	0.4222	0.555
80	0.582	0.521	0.658
100	0.659	0.597	0.731
120	0.719	0.659	0.785

Table 8-15: Probability of DM Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.220	0.186	0.267
40	0.406	0.353	0.477
60	0.538	0.475	0.618
80	0.646	0.581	0.724
100	0.724	0.661	0.796
120	0.784	0.724	0.848

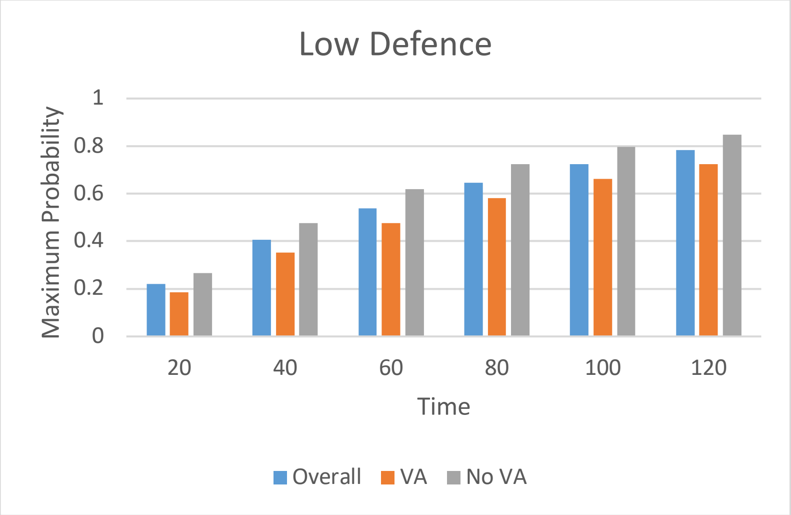


Figure 8-15: Probability of DM Success for Low Defence.

1. When the attacker launches an attack to the components that do not require VA, then it can gain the highest chance for probability of success. This is because application is one of the components that do not require vicinity access, and since it has the least security amongst all the other available component presents, therefore the probability of success is highest for this case. Furthermore, as compared to the previous attack objective, here the attacker requires significantly more time to reach its maximum potential for success regardless of the defence level present. The reason for this behaviour is that it is difficult for the attacker to modify patient data as it requires more effort and requires the attacker to not only access the component but also have write privileges as well.
2. For the case when the attacker has VA, the attacker acquires the least probability of success irrespective of the defence level present. This is because here the attacker can only launch an attack on the sensor, microcontroller, or the gateway. Since all these components have a higher security as compared to the application that is present in the previous case, it is therefore tougher to execute an attack in this case as compared to the case when the attacker has no VA. Furthermore, in this case the difference between all the three defence levels is less compared to the same case for the previous attack objective. This is because here for this attack there does not exist any links present that have encryption mechanisms which play a role in showing some differences between

the high and medium defence level as compared to the low defence level that does not compromise of any encryption mechanism.

3. For the case when the attacker has access to all the components, then the probability of attack success is higher than the VA but lower when the attacker has no access to vicinity. This is because, since the attacker has access to all the components to launch an attack, this means it also comprises of high level of security components which makes it slightly tougher for the attacker as compared to the case when the attacker does not have an option to access these highly secured components.

The results obtained for the cost in terms of the three defence-level can be inferred from Figure 8-16 with Table 8-16, Figure 8-17 with Table 8-17, Figure 8-18 with Table 8-18. Through the references of all these tables and figures, the following observations were made:

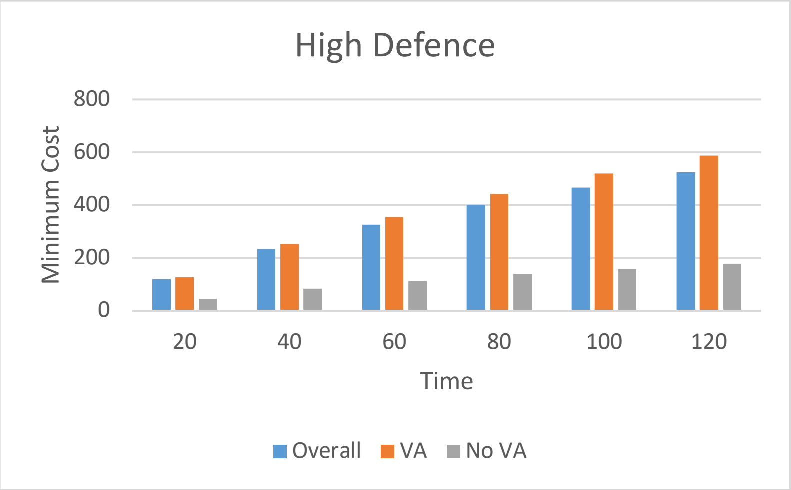


Figure 8-16: DM Cost for High Defence.

Table 8-16: DM Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	119.986	127.082	43.851
40	234.659	252.734	83.679
60	324.807	354.560	113.763
80	400.753	441.386	138.569
100	467.266	520.108	159.451
120	523.601	587.298	177.025

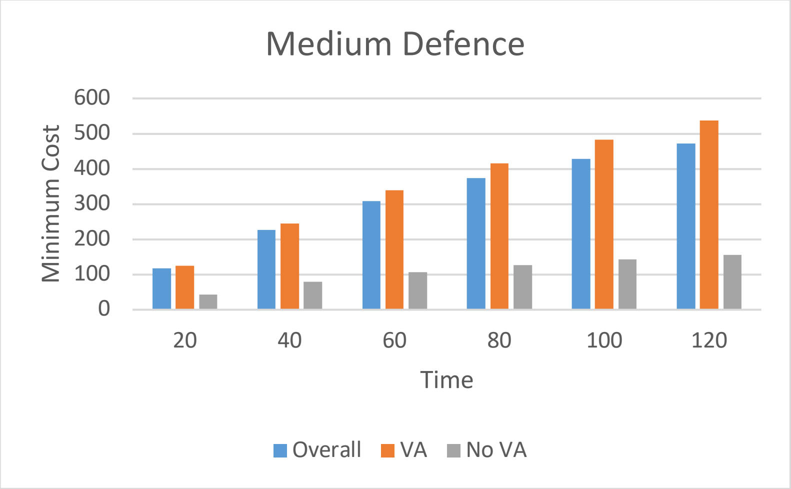


Figure 8-17: DM Cost for Medium Defence.

Table 8-17: DM Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	118.184	125.489	43.031
40	226.832	245.628	80.246
60	308.471	339.585	106.705
80	373.473	415.899	127.079
100	428.148	483.091	143.234
120	471.697	537.685	155.790

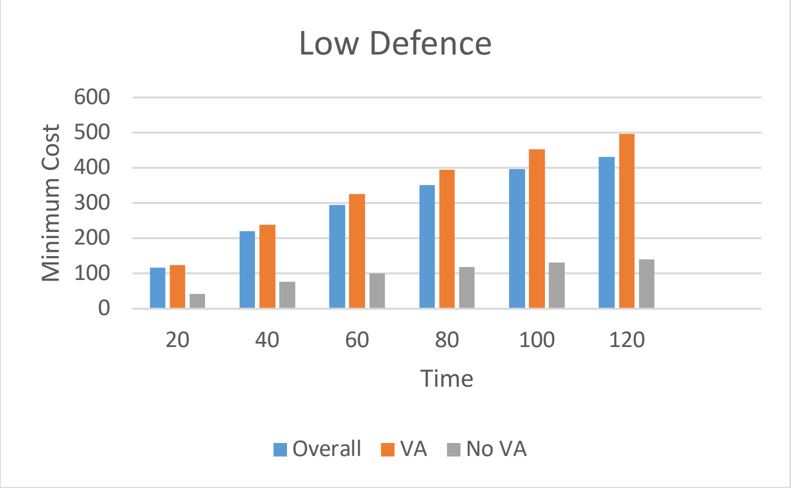


Figure 8-18: DM Cost for Low Defence.

Table 8-18: DM Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	116.539	124.030	42.284
40	219.848	239.242	77.215

60	294.370	326.500	100.722
80	350.561	394.128	117.672
100	396.246	452.223	130.447
120	430.638	497.362	139.664

1. In the case, when the attacker has no vicinity access, the cost for executing the DM attack objective is the least for the same reason explained in the case of the probability of attack success.
2. For the case, when the attacker has access to only VA required components, here since the execution of DM is difficult (because of highly secured components present), therefore the cost of the attack as expected is significantly higher than compared to the case when the attacker has no vicinity access.
3. Finally, in the case when the attacker has access to all the components, here the cost of DM success in this case corresponds to the probability of DM success. That is, the attack cost is higher than compared to the DM cost for when the attacker does not have access to vicinity required components, but it is lesser when the attacker has access to vicinity required components.

For the skilful attacker who chooses the DM as its attack objective, the results obtained for the probability of success in terms of the three defence-level can be inferred from the result generated in Figure 8-19 with Table 8-19, Figure 8-20 with Table 8-20, Figure 8-21 with Table 8-21. Therefore, through the reference of all these figures and tables the following observations were made:

Table 8-19: Probability of DM Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.301	0.135	0.301
40	0.506	0.260	0.506
60	0.633	0.358	0.633
80	0.711	0.445	0.711
100	0.757	0.513	0.757
120	0.784	0.570	0.784

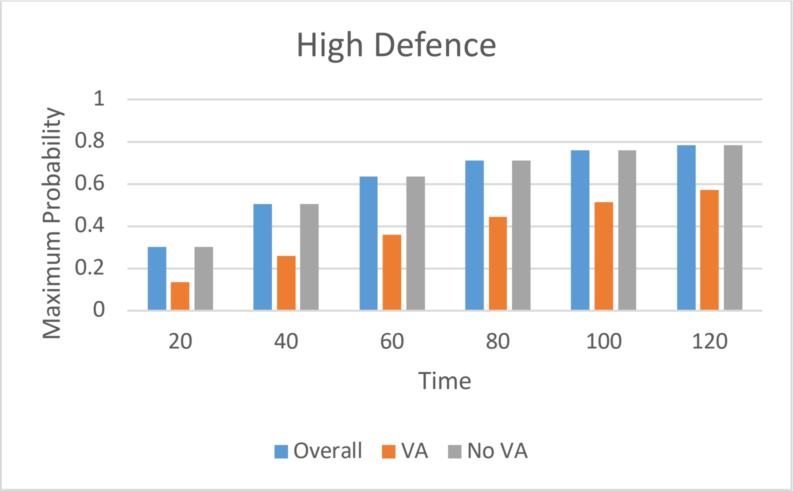


Figure 8-19: Probability of DM Success for High Defence.

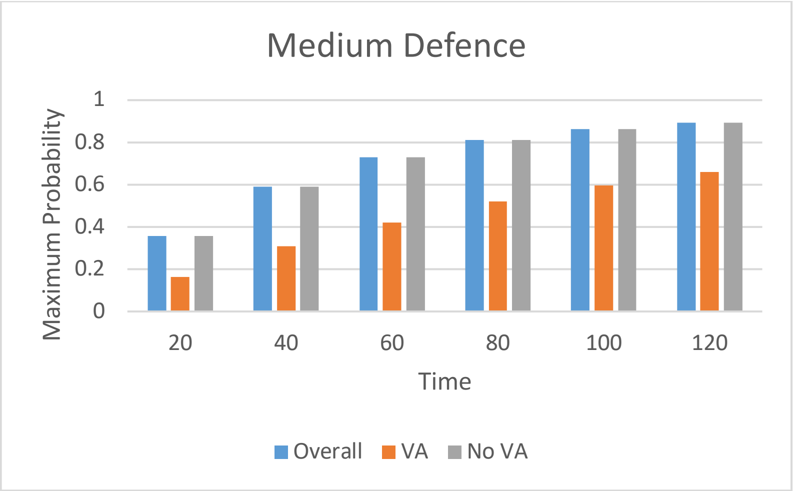


Figure 8-20: Probability of DM Success for Medium Defence.

Table 8-20: Probability of DM Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.358	0.162	0.358
40	0.590	0.310	0.590
60	0.728	0.422	0.728
80	0.812	0.521	0.812
100	0.862	0.597	0.862
120	0.892	0.659	0.892

Table 8-21: Probability of DM Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	0.407	0.186	0.407
40	0.655	0.353	0.655

60	0.794	0.475	0.794
80	0.874	0.581	0.874
100	0.919	0.661	0.919
120	0.946	0.724	0.946

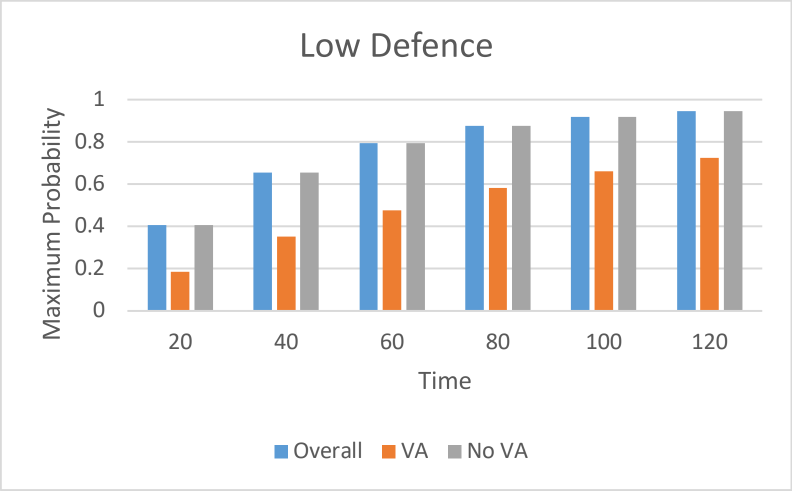


Figure 8-21: Probability of DM Success for Low Defence.

1. The probability of success is same for when the attacker has access to all the components and when the attacker has access to only the components that do not require vicinity access. This is because in this case the attacker is smart enough to choose the component that provides the highest probability success which is the application here. And since for both cases, application is present as a component to launch the DM attack objective, thus the probability of success is same for both.
2. However, there exists a difference when the attacker attacks on components that require VA. This is because the remaining components available have a higher security. Therefore, despite the attacker being smart the probability of attack success would be different and, in this case, worse than the previous case (no VA and access to all components).

The results obtained for the cost in terms of the three defence-level can be inferred from Figure 8-22 with Table 8-22, Figure 8-23 with Table 8-23, Figure 8-24 with Table 8-24. Through the references of all these tables and figures, the following observations were made:

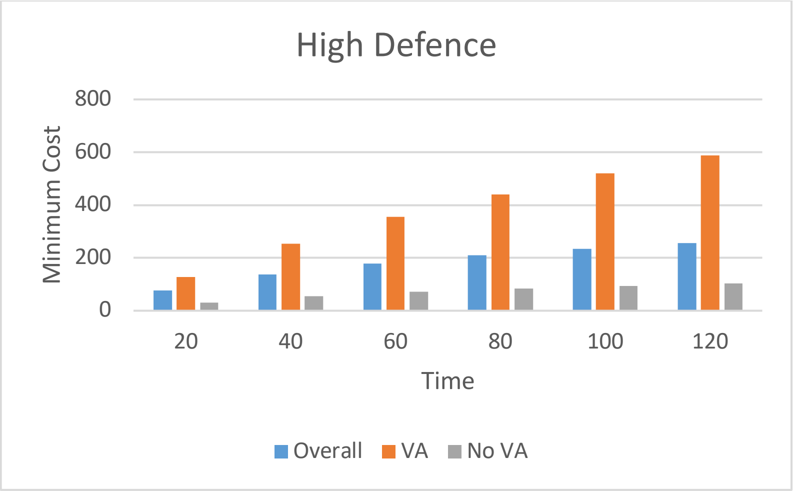


Figure 8-22: DM Cost for High Defence.

Table 8-22: DM Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	77.462	127.082	30.985
40	137.062	252.734	54.824
60	177.945	354.560	71.178
80	209.391	441.386	83.756
100	234.760	520.108	93.904
120	256.585	587.298	102.634

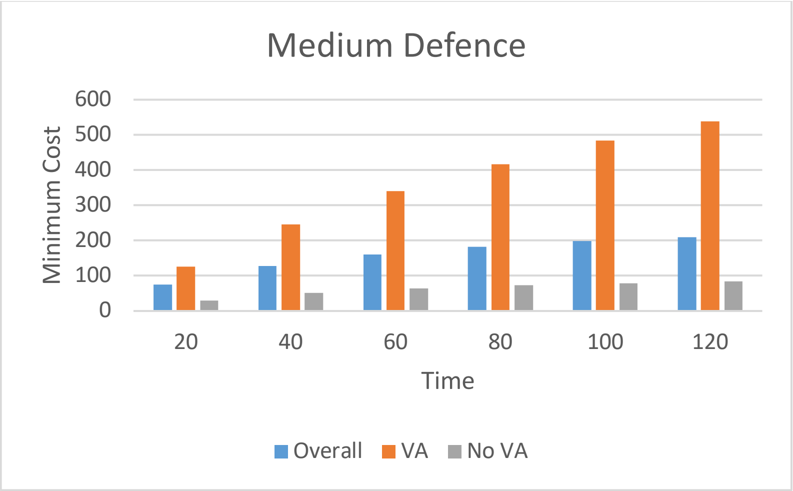


Figure 8-23: DM Cost for Medium Defence.

Table 8-23: DM Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	75.078	125.489	30.031
40	128.096	245.628	51.238

60	160.482	339.585	64.192
80	182.605	415.899	73.042
100	198.227	483.091	79.290
120	209.952	537.685	83.980

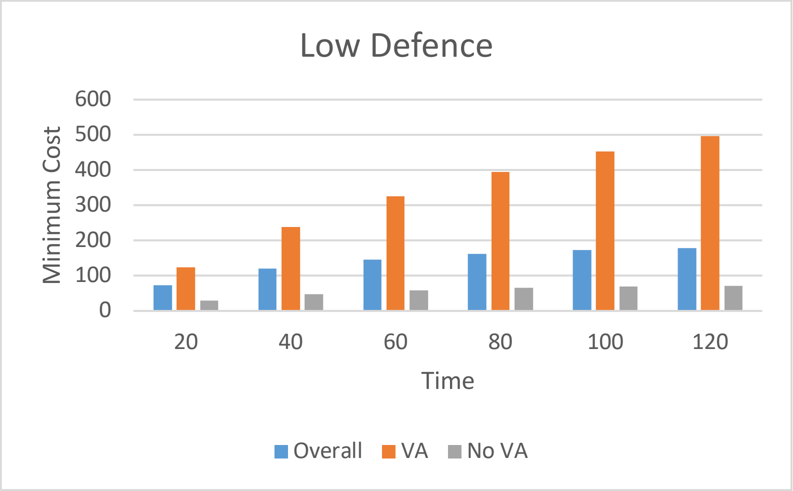


Figure 8-24: DM Cost for Low Defence.

Table 8-24: DM Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
20	72.937	124.030	29.174
40	120.475	239.242	48.190
60	146.551	326.500	58.620
80	162.499	394.128	64.999
100	172.443	452.223	68.977
120	178.895	497.362	71.558

1. The cost for all the three cases differs. This is because for the case when the attacker does not have vicinity access and when the attacker has access to overall the components. Although, the attacker has the option to choose the least secure component (application), however since in general the cost for gaining vicinity access is more, therefore the cost for when the attacker does not have vicinity access is lower than when the attacker has access to all the components.
2. For the case when the attacker has VA, the cost is the highest because of the absence of the application component plus the additional effort require for vicinity access.

8.1.3 Denial-of-Service (DoS)

In the case of launching the DoS attack objective by a naïve attacker, the results obtained for the probability of success in terms of the three defence-level can be inferred from Figure 8-25 with Table 8-25, Figure 8-26 with Table 8-26, Figure 8-27 with Table 8-27. Through the references of all these tables and figures, the following observations were made:

1. For the case when the attacker has no VA, the attacker has a higher probability of attack success as compared to the previous attack objective (DM). However, it is tougher to execute this attack as compared to the DA attack objective. This is because in this attack in addition to gaining access to the component, the attacker needs to shut it down. Furthermore, in terms of probability of success, here the probability of success is the least when the attacker has no VA, because the only options available for this case is $L5$ and data centre, which reduces the chance for the naive attacker to gain a higher probability of success due to the lesser number of links available.
2. For the case when the attacker has access to only VA required components, the probability of success is better than the no VA and is the highest, this is because of the availability of more links ($L1, L2, L3$) apart from $L5$, that increases the chance for the naïve attacker to gain a better probability of success.
3. Lastly, for the case when the attacker has access to all the components, then in this case the attacker reaches the probability of attack success higher when vicinity access is not present, but it is lower than when the vicinity access is present. This is because although it has access to all components, there exits still a possibility of the attacker trying to launch an attack on the $DC.$ However, the presence of other weak components (links) with DC is the reason why in this case the probability of attack success is higher than when the attacker has no vicinity access.

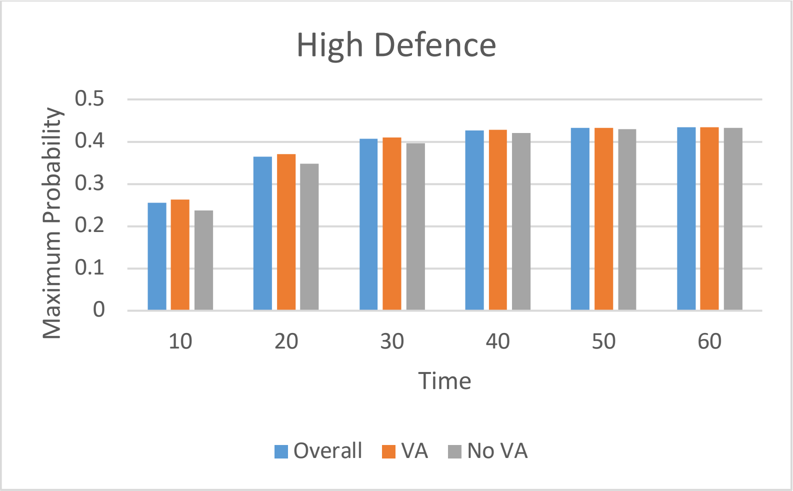


Figure 8-25: Probability of DoS Success for High Defence.

Table 8-25: Probability of DoS Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.255	0.263	0.236
20	0.364	0.370	0.347
30	0.406	0.410	0.396
40	0.426	0.427	0.421
50	0.432	0.432	0.429
60	0.434	0.434	0.432

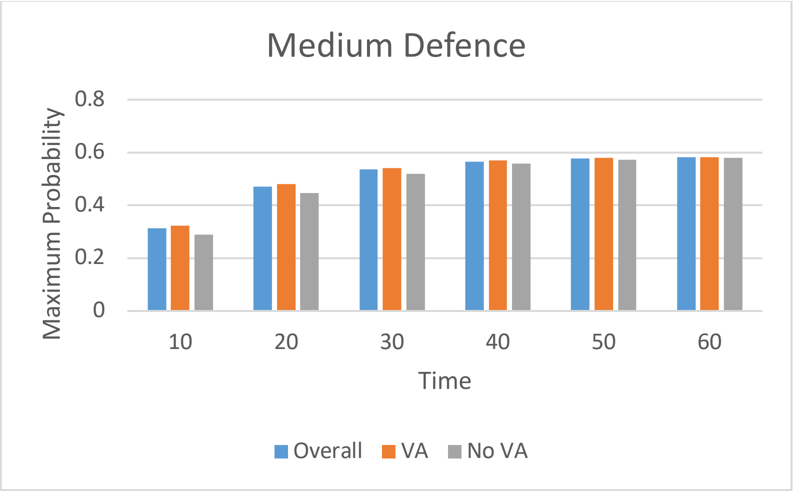


Figure 8-26: Probability of DoS Success for Medium Defence.

Table 8-26: Probability of DoS Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.313	0.322	0.289
20	0.470	0.479	0.446

30	0.535	0.541	0.518
40	0.566	0.569	0.557
50	0.577	0.578	0.572
60	0.581	0.581	0.578

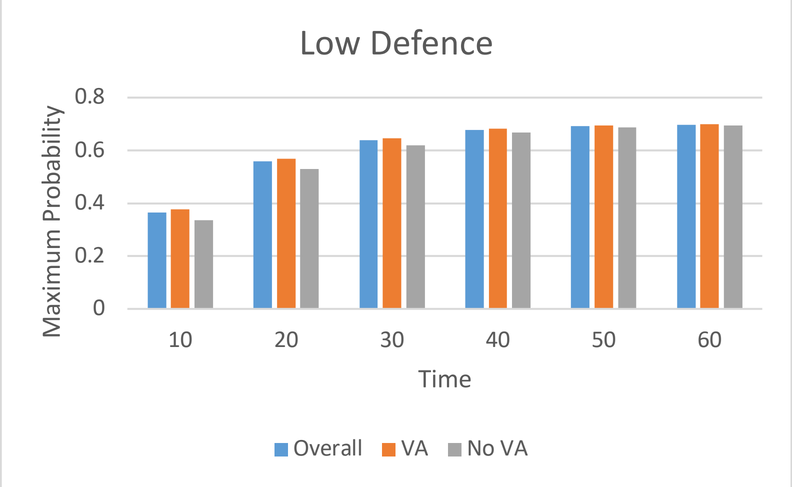


Figure 8-27: Probability of DoS Success for Low Defence.

Table 8-27: Probability of DoS Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.365	0.375	0.336
20	0.558	0.569	0.529
30	0.639	0.646	0.618
40	0.678	0.681	0.666
50	0.693	0.694	0.687
60	0.698	0.698	0.694

The results obtained for the cost in terms of the three defence-level can be inferred from Figure 8-28 with Table 8-28, Figure 8-29 with Table 8-29, Figure 8-30 with Table 8-30. Through the references of all these tables and figures, the following observations were made:

1. On the contrary to its probability success, the cost of the attack when the attacker has no vicinity access is the least. Although for the other cases the attacker has more links as options to launch its attack, but these links besides $L5$ require vicinity access which increases the overall effort required.
2. When the attacker has access to only VA required components, the cost for the execution of the DoS attack, is higher for this as compared to no VA, this is

because of the additional requirement of vicinity access that requires more effort.

3. For the case when the attacker has access to all the components, the cost of DoS is worse than when the attacker has access to only the vicinity component. This is because in this case the attacker has possibility to choose the DC component which increases the attack cost.

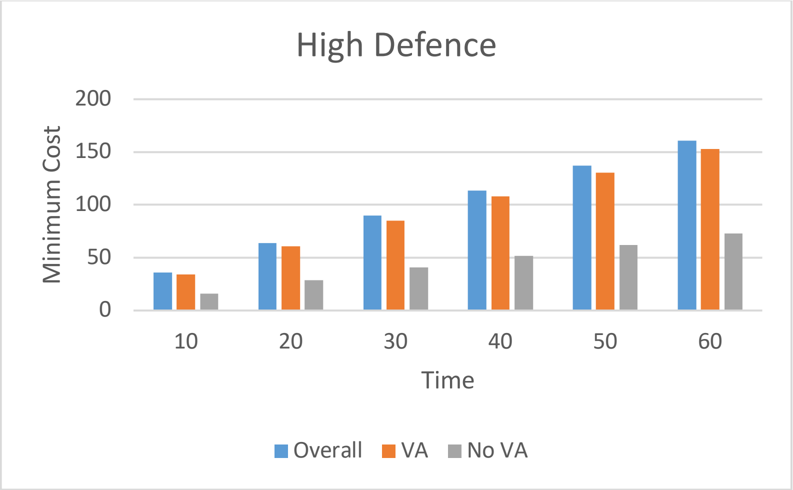


Figure 8-28: DoS Cost for High Defence.

Table 8-28: DoS Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	35.911	34.202	16.102
20	63.916	60.689	28.893
30	89.655	85.014	40.710
40	113.742	107.900	51.613
50	137.378	130.412	62.235
60	160.865	152.806	72.754

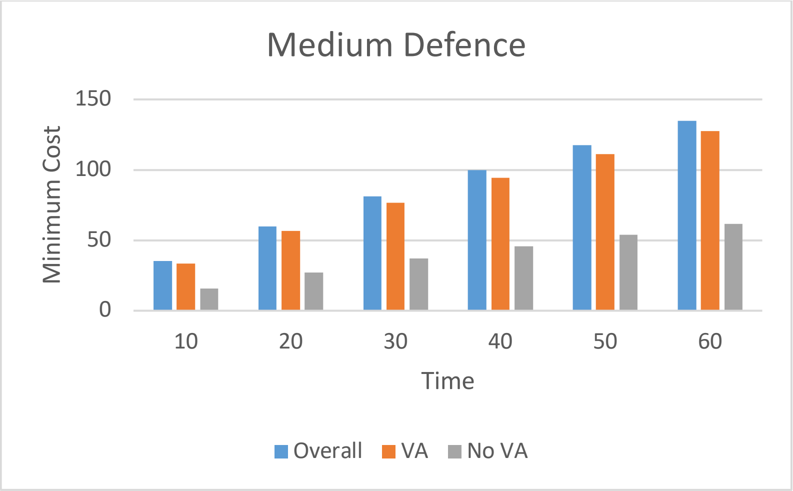


Figure 8-29: DoS Cost for Medium Defence.

Table 8-29: DoS Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	35.135	33.436	15.787
20	59.987	56.815	27.291
30	81.088	76.608	37.174
40	99.678	94.181	45.717
50	117.384	111.000	53.743
60	134.825	127.604	61.597

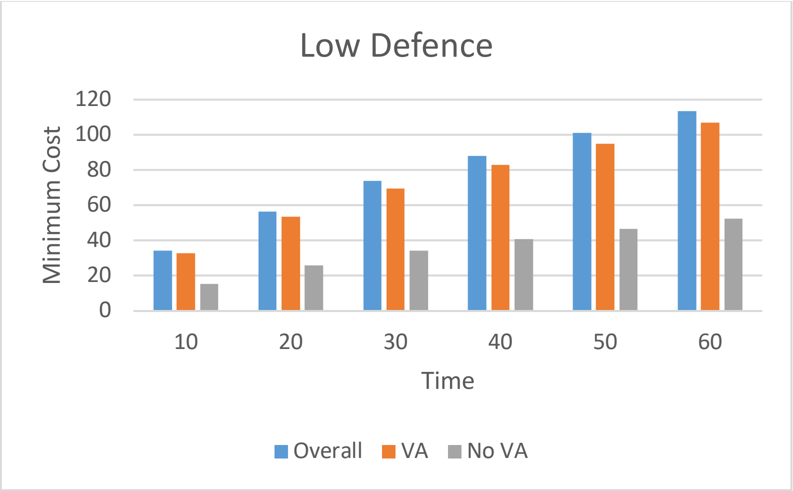


Figure 8-30: DoS Cost for Low Defence.

Table 8-30: DoS Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	34.415	32.724	15.495
20	56.577	53.459	25.893

30	73.874	69.545	34.180
40	88.026	82.835	40.806
50	101.007	95.122	46.758
60	113.645	107.130	52.491

In the case of DoS for a skilful attacker, the results obtained for the probability of success in terms of the three defence-level can be inferred from Figure 8-31 with Table 8-31, Figure 8-32 with Table 8-32, Figure 8-33 with Table 8-33. Through the references of all these tables and figures, the following observations were made:

1. In this case, as the attacker is advanced in terms of skills and experience, therefore, it does not matter whether the attacker has access to the components that require vicinity components, the components that do not require vicinity access, or all the components. This is because in this case there does not exist any encryptions for the links, thus the skilful attacker will always choose the link in all the cases as it is the weakest component. That is, for all the three cases the attacker will reach the same probability of success irrespective of the defence level present.

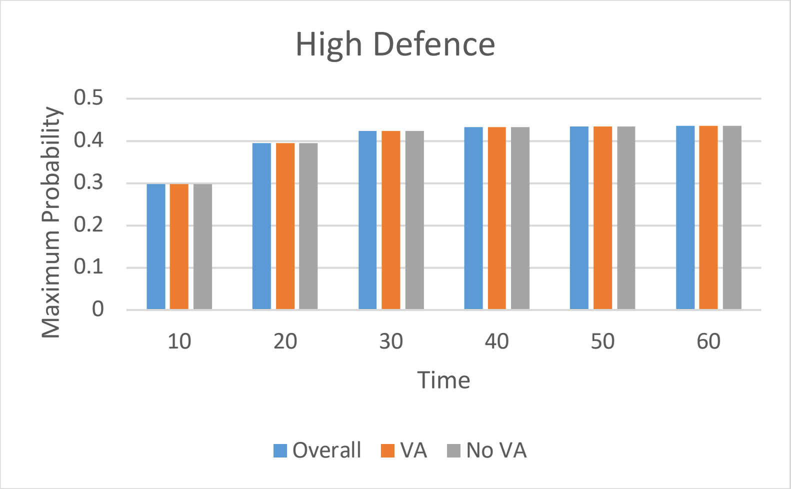


Figure 8-31: Probability of DoS Success for High Defence.

Table 8-31: Probability of DoS Success for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.298	0.298	0.298
20	0.395	0.395	0.395
30	0.423	0.423	0.423
40	0.432	0.432	0.432

50	0.434	0.434	0.434
60	0.435	0.435	0.435

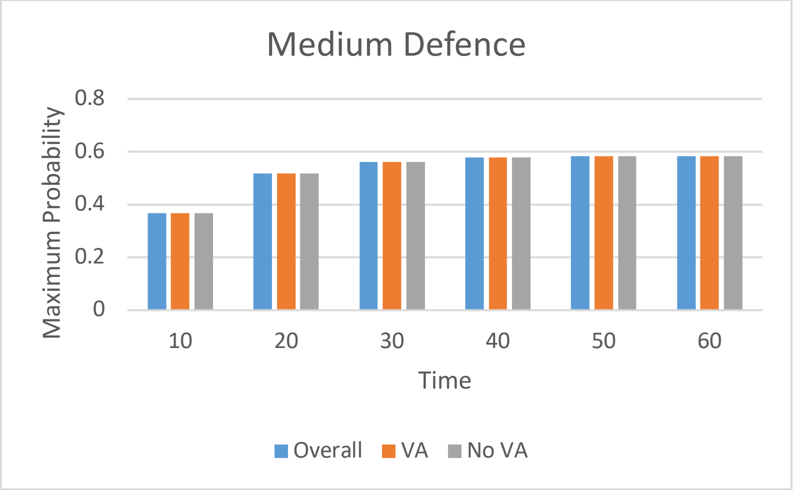


Figure 8-32: Probability of DoS Success for Medium Defence.

Table 8-32: Probability of DoS Success for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.366	0.366	0.366
20	0.516	0.516	0.516
30	0.562	0.562	0.562
40	0.578	0.578	0.578
50	0.582	0.582	0.582
60	0.583	0.583	0.583

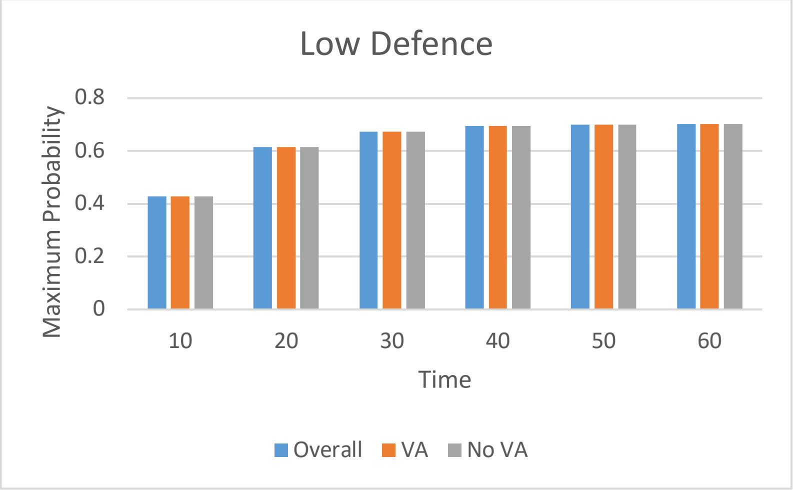


Figure 8-33: Probability of DoS Success for Low Defence.

Table 8-33: Probability of DoS Success for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	0.428	0.428	0.428
20	0.615	0.615	0.615

30	0.673	0.673	0.673
40	0.693	0.693	0.693
50	0.699	0.699	0.699
60	0.700	0.700	0.700

The results obtained for the cost in terms of the three defence-level can be inferred from Figure 8-34 with Table 8-34, Figure 8-35 with Table 8-35, Figure 8-36 with Table 8-36. Through the references of all these tables and figures, the following observations were made:

1. Unlike the probability of success, the cost is different between the case when the attacker does not require VA versus the case when the attacker has access to all the components and only VA required components. This is because the cost for vicinity access in general is higher than without it. Therefore, the cost for when attacker has access to all components and when attacker has access to only vicinity components is the same versus when the attacker does not have access to vicinity components.

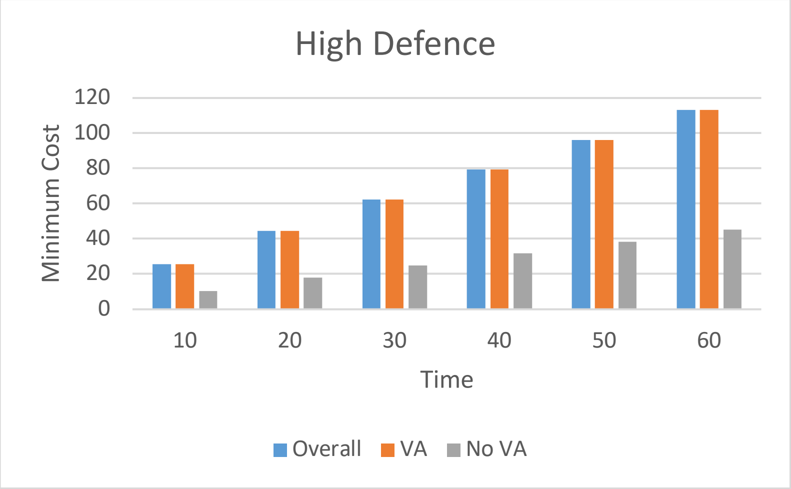


Figure 8-34: DoS Cost for High Defence.

Table 8-34: DoS Cost for High Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	25.506	25.506	10.202
20	44.621	44.621	17.848
30	62.278	62.278	24.911
40	79.297	79.297	31.718
50	96.191	96.191	38.476
60	113.057	113.057	45.222

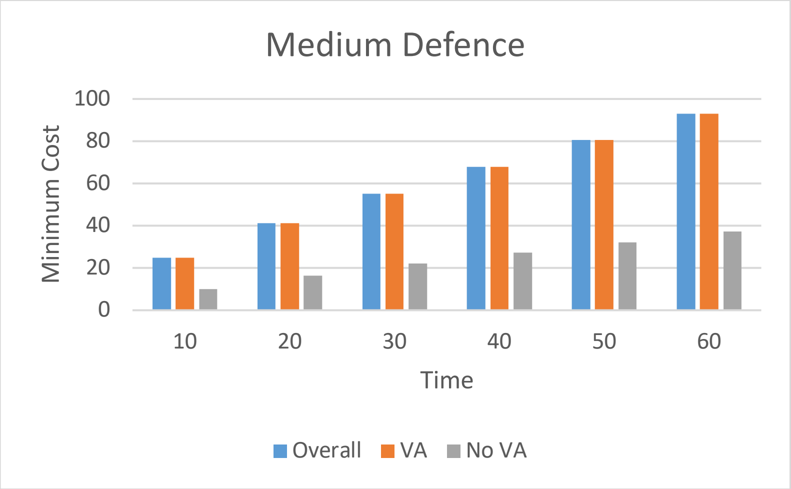


Figure 8-35: DoS Cost for Medium Defence.

Table 8-35: DoS Cost for Medium Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	24.830	24.830	9.932
20	41.239	41.239	16.495
30	55.129	55.129	22.051
40	67.936	67.936	27.174
50	80.451	80.451	32.180
60	92.904	92.904	37.161

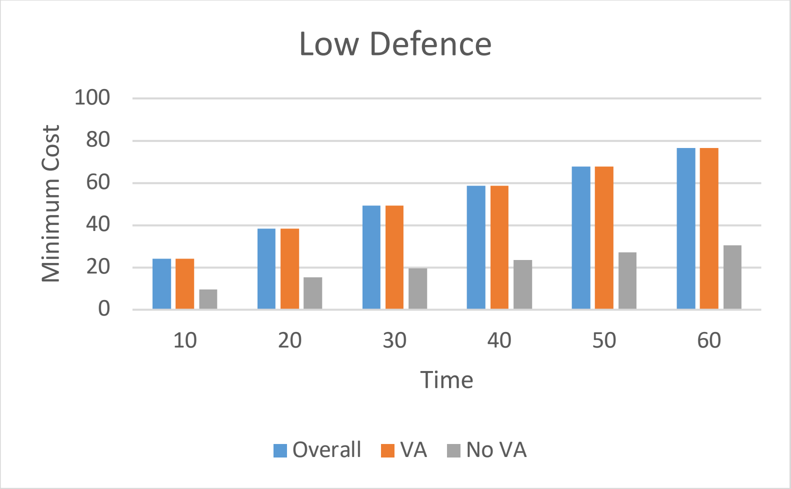


Figure 8-36: DoS Cost for Low Defence.

Table 8-36: DoS Cost for Low Defence.

Time	Overall	Vicinity Access	No Vicinity Access
10	24.202	24.202	9.681
20	38.337	38.337	15.335
30	49.177	49.177	19.670
40	58.618	58.618	23.447
50	67.657	67.657	27.062
60	76.611	76.611	30.644

8.1.4 Ransomware (RW)

Since in this case, the attacker targets only the data centre as its targeted component. Therefore, there is no need to generate the result for when the attacker has access to the vicinity components. Furthermore, there also no need to differentiate between the type of attacker as naive or skilful since there is only one component available as an option for an attacker to execute its attack on.

For the case of probability of attack success, as demonstrated in Figure 8-37 with Table 8-37, the attacker requires a lot of time to increase the probability of attack success. However, this value for probability is still extremely low as compared to the other cases. This is because of the reason that this attack requires to be executed in a series of steps which requires a lot of effort, and time for executing the attack thereby reducing its probability of attack success.

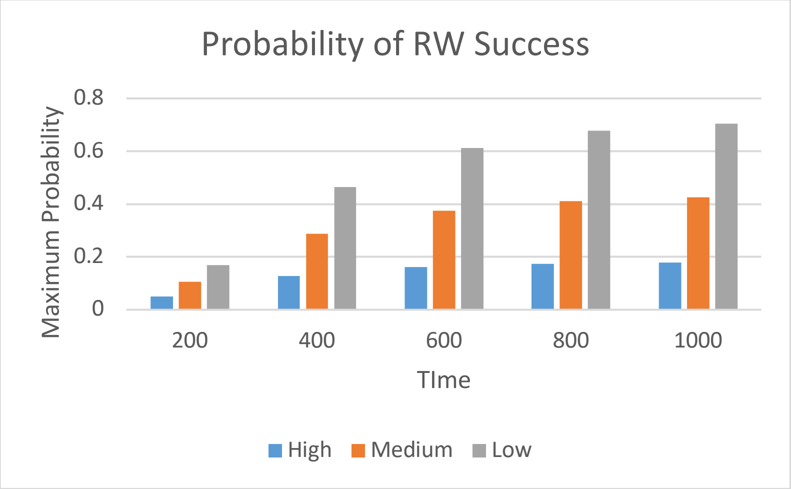


Figure 8-37: Probability of RW Success.

Table 8-37: Probability of RW Success.

Time	High	Medium	Low
200	0.05	0.105	0.169
400	0.128	0.288	0.464
600	0.162	0.375	0.612
800	0.174	0.411	0.677
1000	0.179	0.425	0.704

Corresponding to the probability of success as seen from Figure 8-38 with Table 8-38, the cost of the RW is also extremely high due to the requirement of high effort by the attacker.

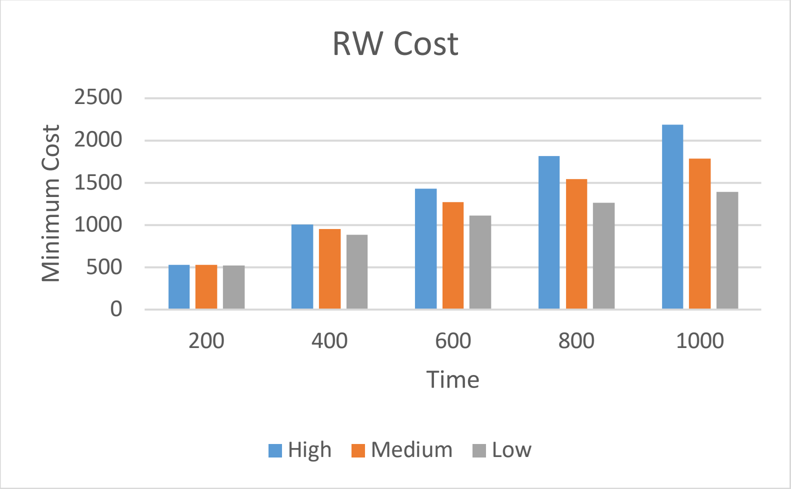


Figure 8-38: RW Cost.

Table 8-38: RW Cost.

Time	High	Medium	Low
200	527.445	526.493	519.950
400	1010.015	953.873	885.666
600	1429.144	1274.843	1110.483
800	1816.953	1542.393	1266.132
1000	2190.655	1784.956	1389.503

8.2. Comparison of Proposed Work with Other Existing Work

Upon the analysis of the related research work regarding security modelling, only the paper published by Mohsin et al. [43] is similar to the proposed work in terms of the security analysis method used. That is, both works utilized PRISM as a tool for security analysis. However, in terms of the technology employed the proposed work utilizes the DT system whereas the other work utilizes the IoT system. Both works have uses cases, however, the proposed work utilized an in-patient monitoring system for the healthcare sector whereas the other work utilizes a home security system, in which they have various configurations for their scenario used that include the addition of the components in their use case, whereas for the use case of the proposed work a constant configuration is used which does not change over time. However, the proposed work can also be incorporated to any kind of configuration since its generic in nature. In terms of defences the compared work does not talk about any defences present in their system, whereas the proposed work has focuses on the defence mechanism for the system. In the defence for the proposed work, two major tier defences, in which the second-tier defence comprises of three defence profile (low, medium, and high). Furthermore, the defences act as soon as the attack is detected by it. In terms of the attack, there exists some differences in the type of attack executed as shown in Table 8-39. Furthermore, the proposed work has attacks that can be executed throughout the system, that is in all the layers present in the system ranging from the physical domain layer to the application layer. Whereas in the compared work, the attacks are only focused on the low level of the system. Apart from this, the attack executed in the proposed work are restricted at a given amount of time, after this it is eventually detected by the defence system. However, in the compared work, there is no mention of time restriction in the attacks executed. Both works, have resource bound attacks, but the proposed work also have an additional type of attacker which is naïve and skilful, whereas the compared work only talks about the skilful attacker. Lastly, the attacks chosen in the proposed work also talk about the different scenarios in the type of attack which is vicinity access, no vicinity access and access to all the components.

For the purpose of quantitative comparison, the probabilities for the attack objectives between the proposed work and an existing approach implemented by Mohsin et al. [43] were compared. Table 8-40 shows a list of these probabilities with their respective attacks that occurs for a time of 100 units where the attacker type is skilful in nature. Furthermore, it is also considered that the attacker has access to all the available components and the defence level considered is high. Here the probability of attack success for all four attacks is significantly higher as compared to the proposed method, this is because the existing approach does not consider a realistic scenario while implementing their model. That is, their model comprises of a negligible defence mechanism which makes it very easy for a skilful attacker in executing its respective attack objective.

Table 8-39: Qualitative Comparison of Existing Work with the Proposed Work.

Features		Mohsin et al.[43]	Proposed Work
Technology Employed		IoT	DT
Application Area Type		Home	Healthcare
Configuration		Variable	Single
Defence Strategy	Levels	N/A	2 Tier Defence Levels
	Profiles	N/A	3 Defence Profile for Tier 2 Defence
	Reaction based on Attack	N/A	Yes
Attack	Type	Theft, missed both alarms, missed either alarm, False alarm,	Data access, data modification, Denial-of-service,

		compromised evacuation	ransomware t
	Level	Low	System
	Time-bound	No	Yes
	Resource bound	Yes	Yes
	Configuration	N/A	VA, No VA, Both
	Attacker	Skillful	Naïve and Skillful

Table 8-40: Quantitative Comparison (lower values better results).

Attack Objective	Mohsin et al. Approach [43]	Proposed Approach
Data Access	1.0	0.81
Data Modification	0.99	0.75
Denial of Service	0.99	0.43
Ransomware	0.15	0.0083

Chapter 9. Conclusion and Future Work

In recent years, DT have started to gain a lot of popularity. However, the first use of DT dates back in the late 1960s where DT was used at the NASA Apollo program for generating a DT of a spacecraft in order to study its behaviour. Later, the official use of DT was first officially coined in 2002 by Michael Grieves who defined DT as a virtual representation of given physical object, process, or system. Typically, the concept of DT is considered to compromise of three major parts which is the physical product, its virtual counterpart and the data and interface connection that exists between them. At present, it has been rapidly employed over a variety of application domains like manufacturing, healthcare, automobile, retail, etc. Nevertheless, the rise in the implementation of DT poses serious concern as there exists a variety of security challenges that can be faced regardless of the industrial domain it is employed in. Therefore, in order to guarantee a smooth execution of the DT operation, a highly secure DT environment is required.

Currently, none of the research work focuses on addressing the security issues faced by the DT system. Therefore, the focus of the thesis is to focus on the security aspect of DT. Before doing so, a modified conceptual model of DT is first presented to understand the working of the DT system. Next, a thorough analysis of the various security attacks and relevant countermeasures were conducted. Based on this analysis, a multi-layer DT security architecture was proposed. However, to implement the proposed security architecture, a probabilistic model checking approach was utilized where DTMC and MDP models were used to represent the behaviour of the naïve and skilful attacker respectively. To provide security to the proposed model, two tier level defences were used, in which tier two defence further comprises of low, medium, and high defence respectively. Furthermore, to validate the proposed approach, a case study of DT in the healthcare sector in the form of in-patient monitoring system was chosen as the proposed scenario. The obtained results after the implementation of the proposed multi-layer DT security architecture demonstrates the maximum probability of success acquired by executing four different attack objectives and the corresponding cost required for their execution. Furthermore, to compare the proposed work with an existing research work, both qualitative and quantitative comparisons were conducted in which the qualitative comparison proved that the proposed work has more advantages than the compared work, whereas the results obtained from the qualitative comparison proved that the proposed work obtained a more realistic result.

In terms of future work, the proposed work can be further expanded in other key areas of DT in the healthcare sector such as out-patient monitoring system. Apart from this, the security aspect of other major areas of DT applications such as manufacturing, energy, and the automobile sector, etc., can also be examined. To further enhance the proposed work a third-tier defence can also be incorporated, where a human operator would be involved whose role is to monitor any disturbances present within the DT system. The addition of this third tier will help in further enhancing the overall security of the DT system.

References

[1] J.-P. A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab, and M. Malli, “Cyber-physical systems security: Limitations, issues and future trends,” *Microprocess. Microsyst.* vol. 77, p. 103201, Sep. 2020, doi: 10.1016/j.micpro.2020.103201.

[2] R. Minerva, G. M. Lee, and N. Crespi, “Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models,” *Proc. IEEE.* vol. 108, no. 10, pp. 1785–1824, Oct. 2020, doi: 10.1109/JPROC.2020.2998530.

[3] M. J. Kaur, V. P. Mishra, and P. Maheshwari, “The Convergence of Digital Twin, IoT, and Machine Learning: Transforming Data into Action,” in *Digital Twin Technologies and Smart Cities*, M. Farsi, A. Daneshkhah, A. Hosseinian-Far, and H. Jahankhani, Eds. Cham: Springer International Publishing, 2020, pp. 3–17. doi: 10.1007/978-3-030-18732-3_1.

[4] S. Bhandari. “Difference Between Digital Twin and IoT (With Table).” https://askanydifference.com/difference-between-digital-twin-and-iot-with-table/ [June. 28, 2022].

[5] M. Grieves, “Digital twin: manufacturing excellence through virtual factory replication." White paper, 2015 https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf [June. 28, 2022].

[6] F. Pires, A. Cachada, J. Barbosa, A. P. Moreira, and P. Leitao, “Digital Twin in Industry 4.0: Technologies, Applications and Challenges,” in *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*. Helsinki, Finland, Jul. 2019, pp. 721–726. doi: 10.1109/INDIN41052.2019.8972134.

[7] R. G. Alves *et al.*, “A digital twin for smart farming,” in *2019 IEEE Global Humanitarian Technology Conference (GHTC)*. Seattle, WA, USA, Oct. 2019, pp. 1–4. doi: 10.1109/GHTC46095.2019.9033075.

[8] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, “Digital Twin in manufacturing: A categorical literature review and classification,” *IFAC-Pap.* vol. 51, no. 11, pp. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.

[9] T. Erol, A. F. Mendi, and D. Dogan, “The Digital Twin Revolution in Healthcare,” in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. Istanbul, Turkey, Oct. 2020, pp. 1–7. doi: 10.1109/ISMSIT50672.2020.9255249.

[10] Y. Zheng, S. Yang, and H. Cheng, “An application framework of digital twin and its case study,” *J. Ambient Intell. Humaniz. Comput.* vol. 10, no. 3, pp. 1141–1153, Mar. 2019, doi: 10.1007/s12652-018-0911-3.

[11] “More information about PRISM.” https://www.prismmodelchecker.org/about .php [June. 28, 2022].

[12] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen, “About The Importance of Autonomy and Digital Twins for the Future of Manufacturing,” *IFAC-Pap.* vol. 48, no. 3, pp. 567–572, 2015, doi: 10.1016/j.ifacol.2015.06.141.

[13] M. Grieves, “Product lifecycle management: the new paradigm for enterprises,” *International Journal of Product Development*. vol. 2, no. 1–2, pp. 71–84, 2005.

[14] M. Grieves, *Product Lifecycle Management: Driving the Next Generation of Lean Thinking The McGraw-Hill Co*. 2005.

[15] M. Grieves, *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management.* 2011.

[16] “Types of digital twins.” https://www.ibm.com/ae-en/topics/what-is-a-digital-twin [June. 28, 2022].

[17] T. Plank, “Digital Twins: The 4 Types and their Characteristics,” 2019. https://tributech.io/blog/the-4-types-of-digital-twins [June. 28, 2022].

[18] S. Nandi and S. B Panikkar, “Building the digital representation with Digital Twin using AWS stack.” https://www.ibm.com/blogs/aws/building-the-digital-representation-with-digital-twin-using-aws-stack/, 2021 [June. 28, 2022].

[19] O. Mbaabu, “Introduction to Digital Twin Technology,” 2020. https://www.section.io/engineering-education/introduction-to-digital-twin-technology/ [June. 28, 2022].

[20] C. Cimino, E. Negri, and L. Fumagalli, “Review of digital twin applications in manufacturing,” *Comput. Ind.* vol. 113, p. 103130, Dec. 2019, doi: 10.1016/j.compind.2019.103130.

[21] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, “Digital twin driven prognostics and health management for complex equipment,” *CIRP Ann.* vol. 67, no. 1, pp. 169–172, 2018, doi: 10.1016/j.cirp.2018.04.055.

[22] Y. Liu *et al.*, “A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin,” *IEEE Access*. vol. 7, pp. 49088–49101, 2019, doi: 10.1109/ACCESS.2019.2909828.

[23] Z. Wang *et al.*, “A Digital Twin Paradigm: Vehicle-to-Cloud Based Advanced Driver Assistance Systems,” in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, May 2020, pp. 1–6. doi: 10.1109/VTC2020-Spring48590.2020.9128938.

[24] “Digital twin in consumer choice modeling,” in *Advances in Computers*. vol. 117, Elsevier, 2020, pp. 265–284.

[25] A. K. Sleiti, J. S. Kapat, and L. Vesely, “Digital twin in energy industry: Proposed robust digital twin for power plant and other complex capital-intensive large engineering systems,” *Energy Rep.* vol. 8, pp. 3704–3726, Nov. 2022, doi: 10.1016/j.egyr.2022.02.305.

[26] B. R. Barricelli, E. Casiraghi, and D. Fogli, “A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications,” *IEEE Access*. vol. 7, pp. 167653–167671, 2019, doi: 10.1109/ACCESS.2019.2953499.

[27] A. Fuller, Z. Fan, C. Day, and C. Barlow, “Digital Twin: Enabling Technologies, Challenges and Open Research,” *IEEE Access*. vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358.

[28] Y. Zheng, S. Yang, and H. Cheng, “An application framework of digital twin and its case study,” *J. Ambient Intell. Humaniz. Comput.* vol. 10, no. 3, pp. 1141–1153, Mar. 2019, doi: 10.1007/s12652-018-0911-3.

[29] J.-F. Uhlenkamp, K. Hribernik, S. Wellsandt, and K.-D. Thoben, “Digital Twin Applications : A first systemization of their dimensions,” in *2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, Valbonne Sophia-Antipolis, France, Jun. 2019, pp. 1–8. doi: 10.1109/ICE.2019.8792579.

[30] Y. Lu, C. Liu, K. I.-K. Wang, H. Huang, and X. Xu, “Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues,” *Robot. Comput.-Integr. Manuf.* vol. 61, p. 101837, Feb. 2020, doi: 10.1016/j.rcim.2019.101837.

[31] J. A. Marmolejo-Saucedo, M. Hurtado-Hernandez, and R. Suarez-Valdes, “Digital Twins in Supply Chain Management: A Brief Literature Review,” in *Intelli-*

gent Computing and Optimization. vol. 1072, P. Vasant, I. Zelinka, and G.-W. Weber, Eds. Cham: Springer International Publishing, 2020, pp. 653–661. doi: 10.1007/978-3-030-33585-4_63.

[32] G. Bhatti, H. Mohan, and R. Raja Singh, "Towards the future of smart electric vehicles: Digital twin technology," *Renew. Sustain. Energy Rev.* vol. 141, p. 110801, May 2021, doi: 10.1016/j.rser.2021.110801.

[33] L. F. Rivera, M. Jiménez, P. Angara, N. M. Villegas, G. Tamura, and H. A. Müller, "Towards Continuous Monitoring in Personalized Healthcare through Digital Twins," p. 6, 2019.

[34] C. Verdouw, B. Tekinerdogan, A. Beulens, and S. Wolfert, "Digital twins in smart farming," *Agric. Syst.* vol. 189, p. 103046, Apr. 2021, doi: 10.1016/j.agsy.2020.103046.

[35] H. Laaki, Y. Miche, and K. Tammi, "Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery," *IEEE Access*. vol. 7, pp. 20325–20336, 2019, doi: 10.1109/ACCESS.2019.2897018.

[36] K. Sivalingam, M. Sepulveda, M. Spring, and P. Davies, "A Review and Methodology Development for Remaining Useful Life Prediction of Offshore Fixed and Floating Wind turbine Power Converter with Digital Twin Technology Perspective," in *2018 2nd International Conference on Green Energy and Applications (ICGEA)*, Singapore, Mar. 2018, pp. 197–204. doi: 10.1109/ICGEA.2018.8356292.

[37] X. Chen, E. Kang, S. Shiraishi, V. M. Preciado, and Z. Jiang, "Digital Behavioral Twins for Safe Connected Cars," in *Proceedings of the 21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, Copenhagen Denmark, Oct. 2018, pp. 144–153. doi: 10.1145/3239372.3239401.

[38] A. R. Al-Ali, R. Gupta, T. Zaman Batool, T. Landolsi, F. Aloul, and A. Al Nabulsi, "Digital Twin Conceptual Model within the Context of Internet of Things," *Future Internet*. vol. 12, no. 10, p. 163, Sep. 2020, doi: 10.3390/fi12100163.

[39] G. Potrino, F. de Rango, and A. F. Santamaria, "Modeling and evaluation of a new IoT security system for mitigating DoS attacks to the MQTT broker," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, Marrakesh, Morocco, Apr. 2019, pp. 1–6. doi: 10.1109/WCNC.2019.8885553.

[40] M. Ge, J. B. Hong, W. Guttmann, and D. S. Kim, "A framework for automating security analysis of the internet of things," *J. Netw. Comput. Appl.* vol. 83, pp. 12–27, Apr. 2017, doi: 10.1016/j.jnca.2017.01.033.

[41] P. K. Dhillon and S. Kalra, "A lightweight biometrics based remote user authentication scheme for IoT services," *J. Inf. Secur. Appl.* vol. 34, pp. 255–270, Jun. 2017, doi: 10.1016/j.jisa.2017.01.003.

[42] S. Zahra *et al.*, "Fog Computing Over IoT: A Secure Deployment and Formal Verification," *IEEE Access*. vol. 5, pp. 27132–27144, 2017, doi: 10.1109/ACCESS.2017.2766180.

[43] M. Mohsin, M. U. Sardar, O. Hasan, and Z. Anwar, "IoTRiskAnalyzer: A Probabilistic Model Checking Based Framework for Formal Risk Analytics of the Internet of Things," *IEEE Access*. vol. 5, pp. 5494–5505, 2017, doi: 10.1109/ACCESS.2017.2696031.

[44] P. K. Dhillon and S. Kalra, "Multi-factor user authentication scheme for IoT-based healthcare services," *J. Reliab. Intell. Environ.* vol. 4, no. 3, pp. 141–160, Sep. 2018, doi: 10.1007/s40860-018-0062-5.

[45] B. D. Deebak, F. Al-Turjman, M. Aloqaily, and O. Alfandi, “An Authentic-Based Privacy Preservation Protocol for Smart e-Healthcare Systems in IoT,” *IEEE Access*. vol. 7, pp. 135632–135649, 2019, doi: 10.1109/ACCESS.2019.2941575.

[46] F. Merabet, A. Cherif, M. Belkadi, O. Blazy, E. Conchon, and D. Sauveron, “New efficient M2C and M2M mutual authentication protocols for IoT-based healthcare applications,” *Peer--Peer Netw. Appl.* vol. 13, no. 2, pp. 439–474, Mar. 2020, doi: 10.1007/s12083-019-00782-8.

[47] N. Theera-Umpon, K.-H. Han, W.-S. Bae, and S. Lee, “Verifying Secure Authentication Protocol for Communication between IoT-based Medical Devices,” p. 13.

[48] A. Rasheed, O. San, and T. Kvamsdal, “Digital Twin: Values, Challenges and Enablers From a Modeling Perspective,” *IEEE Access*. vol. 8, pp. 21980–22012, 2020, doi: 10.1109/ACCESS.2020.2970143.

[49] A. Madni, C. Madni, and S. Lucero, “Leveraging Digital Twin Technology in Model-Based Systems Engineering,” *Systems*. vol. 7, no. 1, p. 7, Jan. 2019, doi: 10.3390/systems7010007.

[50] P. P. Ray, D. Dash, and N. Kumar, “Sensors for internet of medical things: State-of-the-art, security and privacy issues, challenges and future directions,” *Comput. Commun.* vol. 160, pp. 111–131, Jul. 2020, doi: 10.1016/j.comcom.2020.05.029.

[51] "Apple Watch. Helping your patients identify early warning signs." https://www.apple.com/ae/healthcare/apple-watch/, [June. 28, 2022].

[52] "Galaxy Watch4 Classic." https://www.samsung.com/us/watches/galaxy-watch4-classic/#health, [June. 28, 2022].

[53] J. Su, Z. Sheng, V. C. M. Leung, and Y. Chen, “Energy Efficient Tag Identification Algorithms For RFID: Survey, Motivation And New Design,” *IEEE Wirel. Commun.* vol. 26, no. 3, pp. 118–124, Jun. 2019, doi: 10.1109/MWC.2019.1800249.

[54] J. Janata, *Principles of chemical sensors*, 2nd ed. Dordrecht ; New York: Springer, 2009.

[55] "Raspberry Pi 4 Tech Specs." https://www.raspberrypi.org/products/raspberry-pi-4-model-b/specifications/, [June. 28, 2022].

[56] "Particle Photon: Wi-Fi." https://docs.particle.io/photon/, [June. 28, 2022].

[57] "ESP32." https://www.espressif.com/en/products/socs/esp32, [June. 28, 2022].

[58] "IoT Ethernet Kit." https://www.microchip.com/en-us/development-tool/DM990004, [June. 28, 2022].

[59] "Intel Quark Microcontroller D2000." https://www.intel.com/content/www/us/en/products/sku/91947/, [June. 28, 2022].

[60] A. Gupta and R. K. Jha, “A Survey of 5G Network: Architecture and Emerging Technologies,” *IEEE Access*. vol. 3, pp. 1206–1232, 2015, doi: 10.1109/ACCESS.2015.2461602.

[61] M. Alsabah *et al.*, “6G Wireless Communications Networks: A Comprehensive Survey,” *IEEE Access*. vol. 9, pp. 148191–148243, 2021, doi: 10.1109/ACCESS.2021.3124812.

[62] "Digital Twin Market." https://www.marketsandmarkets.com/Market-Reports/digital-twin-market-225269522.html, [June. 28, 2022].

[63] A. K. Singh and B. D. K. Patro, “Security Attacks on RFID and their Countermeasures,” in *Computer Communication, Networking and IoT*, Springer Singapore, 2021, pp. 509–518.

[64] A. Kamble and S. Bhutad, “Survey on Internet of Things (IoT) security issues & solutions,” in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, Jan. 2018, pp. 307–312. doi: 10.1109/ICISC.2018.8399084.

[65] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, “A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications,” *IEEE Internet Things J.* vol. 4, no. 5, pp. 1125–1142, Oct. 2017, doi: 10.1109/JIOT.2017.2683200.

[66] C. Gao, L. Luo, Y. Zhang, B. Pearson, and X. Fu, “Microcontroller Based IoT System Firmware Security: Case Studies,” in *2019 IEEE International Conference on Industrial Internet (ICII)*, Orlando, FL, USA, Nov. 2019, pp. 200–209. doi: 10.1109/ICII.2019.00045.

[67] Y. Shah and S. Sengupta, “A survey on Classification of Cyber-attacks on IoT and IIoT devices,” in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, Oct. 2020, pp. 0406–0413. doi: 10.1109/UEMCON51285.2020.9298138.

[68] G. Palavicini, J. Bryan, E. Sheets, M. Kline, and J. Miguel, “Towards Firmware Analysis of Industrial Internet of Things (IIoT) - Applying Symbolic Analysis to IIoT Firmware Vetting,” 2017. doi: 10.5220/0006393704700477.

[69] D. Papp, Z. Ma, and L. Buttyan, “Embedded systems security: Threats, vulnerabilities, and attack taxonomy,” in *2015 13th Annual Conference on Privacy, Security and Trust (PST)*, Izmir, Turkey, Jul. 2015, pp. 145–152. doi: 10.1109/PST.2015.7232966.

[70] D. Papp, Z. Ma, and L. Buttyan, “Embedded systems security: Threats, vulnerabilities, and attack taxonomy,” in *2015 13th Annual Conference on Privacy, Security and Trust (PST)*, Jul. 2015, pp. 145–152. doi: 10.1109/PST.2015.7232966.

[71] R. Nair, P. Sharma, and A. Khamparia, “Security Attacks in Internet of Things,” in *Fog, Edge, and Pervasive Computing in Intelligent IoT Driven Applications*, 2020, pp. 237–261.

[72] N. Abughazaleh, R. Bin, M. Btish, and H. M., “DoS Attacks in IoT Systems and Proposed Solutions,” *Int. J. Comput. Appl.* vol. 176, no. 33, pp. 16–19, Jun. 2020, doi: 10.5120/ijca2020920397.

[73] M. A. Iqbal and O. G. Olaleye, “A Review on Internet of Things (Iot): Security and Privacy Requirements and the Solution Approaches,” p. 11, 2016.

[74] H. Teymourlouei, “Quick Reference: Cyber Attacks Awareness and Prevention Method for Home Users,” *International Journal of Computer and Systems Engineering.* vol. 9, no. 3, pp. 678-684, 2015, Accessed: Mar. 29, 2022.

[75] J. Granjal, E. Monteiro, and J. Sá Silva, “Security for the Internet of Things: A Survey of Existing Protocols and Open Research Issues,” *IEEE Commun. Surv. Tutor.* vol. 17, no. 3, pp. 1294–1312, 2015, doi: 10.1109/COMST.2015.2388550.

[76] G. Rajendran, R. S. Ragul Nivash, P. P. Parthy, and S. Balamurugan, “Modern security threats in the Internet of Things (IoT): Attacks and Countermeasures,” in *2019 International Carnahan Conference on Security Technology (ICCST)*, CHENNAI, India, Oct. 2019, pp. 1–6. doi: 10.1109/CCST.2019.8888399.

[77] H. A. Abdulghani, N. A. Nijdam, A. Collen, and D. Konstantas, “A Study on Security and Privacy Guidelines, Countermeasures, Threats: IoT Data at Rest Perspective,” *Symmetry*. vol. 11, no. 6, p. 774, Jun. 2019, doi: 10.3390/sym11060774.

[78] G. A. Abdalrahman and H. Varol, “Defending Against Cyber-Attacks on the Internet of Things,” in *2019 7th International Symposium on Digital Forensics and Security (ISDFS)*, Barcelos, Portugal, Jun. 2019, pp. 1–6. doi: 10.1109/ISDFS.2019.8757478.

[79] S. Dey and S. K. Sen, “Four Dimensional Security and Vulnerability Matrix for Cloud (4-SVM),” Jun. 2017, pp. 165–169. doi: 10.15439/2017R41.

[80] Ö. Aslan, M. Ozkan-Okay, and D. Gupta, “A Review of Cloud-Based Malware Detection System: Opportunities, Advances and Challenges,” *Eur. J. Eng. Technol. Res.* vol. 6, no. 3, pp. 1–8, Mar. 2021, doi: 10.24018/ejers.2021.6.3.2372.

[81] N. Alomar, M. Alsaleh, and A. Alarifi, “Social Authentication Applications, Attacks, Defense Strategies and Future Research Directions: A Systematic Review,” *IEEE Commun. Surv. Tutor.* vol. 19, no. 2, pp. 1080–1111, 2017, doi: 10.1109/COMST.2017.2651741.

[82] Z. Brakerski and V. Vaikuntanathan, “Efficient Fully Homomorphic Encryption from (Standard) LWE,” in *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, Palm Springs, CA, USA, Oct. 2011, pp. 97–106. doi: 10.1109/FOCS.2011.12.

[83] A. Ait, N. Ammari, A. Abou, A. Ait, and M. De, “New mechanism for Cloud Computing Storage Security,” *Int. J. Adv. Comput. Sci. Appl.* vol. 7, no. 7, 2016, doi: 10.14569/IJACSA.2016.070773.

[84] M. R. Islam and K. M. Aktheruzzaman, “An Analysis of Cybersecurity Attacks against Internet of Things and Security Solutions,” *J. Comput. Commun.* vol. 08, no. 04, pp. 11–25, 2020, doi: 10.4236/jcc.2020.84002.

[85] B. B. Gupta, N. A. G. Arachchilage, and K. E. Psannis, “Defending against phishing attacks: taxonomy of methods, current issues and future directions,” *Telecommun. Syst.* vol. 67, no. 2, pp. 247–267, Feb. 2018, doi: 10.1007/s11235-017-0334-z.

[86] J. Sengupta, S. Ruj, and S. Das Bit, “A Comprehensive Survey on Attacks, Security Issues and Blockchain Solutions for IoT and IIoT,” *J. Netw. Comput. Appl.* vol. 149, p. 102481, Jan. 2020, doi: 10.1016/j.jnca.2019.102481.

[87] A. Shaikh, A. A. Khan, S. Zebanaaz, S. Shaikh, and N. Akhter, “Exploring Recent Challenges in Cyber Security and their Solutions,”*International Journal of Creative Research Thoughts.* vol. 9, no. 12, p. 6, Dec. 2021.

[88] S. N. Swamy, D. Jadhav, and N. Kulkarni, “Security Threats in the Application layer in IOT Applications,” p. 4, 2017.

[89] M. Kwiatkowska, G. Norman, and D. Parker, “Stochastic Model Checking,” p. 50.

[90] S. Arzaghi, “An Introduction To Linear Temporal Logic,” p. 9.

[91] R. Alur and T. A. Henzinger, “Reactive modules,” in *Proceedings 11th Annual IEEE Symposium on Logic in Computer Science*, New Brunswick, NJ, USA, 1996, pp. 207–218. doi: 10.1109/LICS.1996.561320.

[92] "Digital Twin?" https://www.gehccommandcenter.com/digital-twin, [June. 28, 2022].

[93] C. Rustici. "Dassault Systemes Bets on Healthcare and Life Sciences." http://emag.medicalexpo.com/dassault-systemes-bets-on-healthcare-and-life-sciences/, [June. 28, 2022].

[94] "The rise of the digital twin: how healthcare can benefit.″ https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/20180830-the-rise-of-the-digital-twin-how-healthcare-can-benefit.html, [June. 28, 2022].

[95] Y. Fathy and P. Barnaghi, "Quality-Based and Energy-Efficient Data Communication for the Internet of Things Networks," *IEEE Internet Things J.* vol. 6, no. 6, pp. 10318–10331, Dec. 2019, doi: 10.1109/JIOT.2019.2938101.

[96] H. A. El Zouka and M. M. Hosni, "Secure IoT communications for smart healthcare monitoring system," *Internet Things*. vol. 13, p. 100036, Mar. 2021, doi: 10.1016/j.iot.2019.01.003.

[97] R. Kunnavil, "Healthcare Data Utilization for the Betterment of Mankind - An Overview of Big Data Concept in Healthcare," *Int. J. Healthc. Educ. Med. Inform.*, 2018, doi: 10.24321/2455.9199.201807.

[98] V. Dutt and A. Kaur, "Cyber security: testing the effects of attack strategy, similarity, and experience on cyber attack detection," *Int. J. Trust Manag. Comput. Commun.* vol. 1, no. 3/4, p. 261, 2013, doi: 10.1504/IJTMCC.2013.056428.

[99] K. Akomea-Agyin and M. Asante, "Analysis of Security Vulnerabilities in Wired Equivalent Privacy (WEP)," vol. 06, no. 01, p. 8, 2019.

[100] G. Manley. "Why You Shouldn't Use WEP Encryption." https://www.section.io/engineering-education/wep-encryption/ [June. 28, 2022].

[101] Wade Trappe and Lawrence C. Washington, *Introduction to Cryptography with Coding Theory, 3rd Edition*. Pearson, 2020.

[102] A. J. Alhasan and N. Surantha, "Evaluation of Data Center Network Security based on Next-Generation Firewall," *Int. J. Adv. Comput. Sci. Appl.* vol. 12, no. 9, 2021, doi: 10.14569/IJACSA.2021.0120958.

[103] Subhash V. Pingale and Sanjay R. Sutar, "Analysis of Web Application Firewalls, Challenges, and Research Opportunities," in *Lecture Notes in Electrical Engineering*. vol. 783, Singapore: Springer, 2022, pp. 239–248.

[104] M. Nawir, A. Amir, N. Yaakob, and O. B. Lynn, "Internet of Things (IoT): Taxonomy of security attacks," in *2016 3rd International Conference on Electronic Design (ICED)*, Phuket, Thailand, Aug. 2016, pp. 321–326. doi: 10.1109/ICED.2016.7804660.

[105] S. A. Butt, J. L. Diaz-Martinez, T. Jamal, A. Ali, E. De-La-Hoz-Franco, and M. Shoaib, "IoT Smart Health Security Threats," in *2019 19th International Conference on Computational Science and Its Applications (ICCSA)*, Saint Petersburg, Russia, Jul. 2019, pp. 26–31. doi: 10.1109/ICCSA.2019.000-8.

[106] I. Butun, P. Osterberg, and H. Song, "Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures," *IEEE Commun. Surv. Tutor.* vol. 22, no. 1, pp. 616–644, 2020, doi: 10.1109/COMST.2019.2953364.

[107] F. Alkhudhayr, S. Alfarraj, B. Aljameeli, and S. Elkhdiri, "Information Security:A Review of Information Security Issues and Techniques," in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, May 2019, pp. 1–6. doi: 10.1109/CAIS.2019.8769504.

[108] M. Z. Gunduz and R. Das, "Cyber-security on smart grid: Threats and potential solutions," *Comput. Netw.* vol. 169, p. 107094, Mar. 2020, doi: 10.1016/j.comnet.2019.107094.

[109] J. Zhang, H. Chen, L. Gong, J. Cao, and Z. Gu, "The Current Research of IoT Security," in *2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC)*, Hangzhou, China, Jun. 2019, pp. 346–353. doi: 10.1109/DSC.2019.00059.

[110] E. Staddon, V. Loscri, and N. Mitton, "Attack Categorisation for IoT Applications in Critical Infrastructures, a Survey," *Appl. Sci.* vol. 11, no. 16, p. 7228, Aug. 2021, doi: 10.3390/app11167228.

[111] C. Bradley, S. El-Tawab, and M. H. Heydari, "Security analysis of an IoT system used for indoor localization in healthcare facilities," in *2018 Systems and Information Engineering Design Symposium (SIEDS)*, Charlottesville, VA, Apr. 2018, pp. 147–152. doi: 10.1109/SIEDS.2018.8374726.

[112] S. H. Haji and S. Y. Ameen, "Attack and Anomaly Detection in IoT Networks using Machine Learning Techniques: A Review," *Asian J. Res. Comput. Sci.*, pp. 30–46, Jun. 2021, doi: 10.9734/ajrcos/2021/v9i230218.

[113] E. Džaferović, A. Sokol, A. A. Almisreb, and S. Mohd Norzeli, "DoS and DDoS vulnerability of IoT: A review," *Sustain. Eng. Innov.* vol. 1, no. 1, pp. 43–48, Jun. 2019, doi: 10.37868/sei.v1i1.36.

[114] R. Khader and D. Eleyan, "Survey of DoS/DDoS attacks in IoT," *Sustain. Eng. Innov.* vol. 3, no. 1, pp. 23–28, Jan. 2021, doi: 10.37868/sei.v3i1.124.

[115] L. Liang, K. Zheng, Q. Sheng, and X. Huang, "A Denial of Service Attack Method for an IoT System," in *2016 8th International Conference on Information Technology in Medicine and Education (ITME)*, Fuzhou, China, Dec. 2016, pp. 360–364. doi: 10.1109/ITME.2016.0087.

[116] A. Abdullah, R. Hamad, M. Abdulrahman, H. Moala, and S. Elkhediri, "CyberSecurity: A Review of Internet of Things (IoT) Security Issues, Challenges and Techniques," in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, Riyadh, Saudi Arabia, May 2019, pp. 1–6. doi: 10.1109/CAIS.2019.8769560.

[117] A. Raghuprasad, S. Padmanabhan, M. Arjun Babu, and P. K. Binu, "Security Analysis and Prevention of Attacks on IoT Devices," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, Jul. 2020, pp. 0876–0880. doi: 10.1109/ICCSP48568.2020.9182055.

[118] M. Humayun, N. Jhanjhi, A. Alsayat, and V. Ponnusamy, "Internet of things and ransomware: Evolution, mitigation and prevention," *Egypt. Inform. J.* vol. 22, no. 1, pp. 105–117, Mar. 2021, doi: 10.1016/j.eij.2020.05.003.

[119] A. Zahra and M. A. Shah, "IoT based ransomware growth rate evaluation and detection using command and control blacklisting," in *2017 23rd International Conference on Automation and Computing (ICAC)*, Huddersfield, United Kingdom, Sep. 2017, pp. 1–6. doi: 10.23919/IConAC.2017.8082013.

[120] S. R. Zahra and M. Ahsan Chishti, "RansomWare and Internet of Things: A New Security Nightmare," in *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, Jan. 2019, pp. 551–555. doi: 10.1109/CONFLUENCE.2019.8776926.

[121] I. Yaqoob *et al.*, "The rise of ransomware and emerging security challenges in the Internet of Things," *Comput. Netw.* vol. 129, pp. 444–458, Dec. 2017, doi: 10.1016/j.comnet.2017.09.003.

[122] P. O'Kane, S. Sezer, and D. Carlin, "Evolution of ransomware," *IET Netw.* vol. 7, no. 5, pp. 321–327, Sep. 2018, doi: 10.1049/iet-net.2017.0207.

[123] M. Abdur, S. Habib, M. Ali, and S. Ullah, "Security Issues in the Internet of Things (IoT): A Comprehensive Study," *Int. J. Adv. Comput. Sci. Appl.* vol. 8, no. 6, 2017, doi: 10.14569/IJACSA.2017.080650.

[124] A. I. Newaz, A. K. Sikder, M. A. Rahman, and A. S. Uluagac, "A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses," *ACM Trans. Comput. Healthc.* vol. 2, no. 3, pp. 1–44, Jul. 2021, doi: 10.1145/3453176.

[125] R. Chawngsangpuii, P. Das, and R. K. Das, "Security Management perspective for Internet of Things," *Int. J. Eng. Sci. Invent.* vol. 6, no. 9, pp. 59–65.

[126] T. Dargahi, A. Dehghantanha, P. N. Bahrami, M. Conti, G. Bianchi, and L. Benedetto, "A Cyber-Kill-Chain based taxonomy of crypto-ransomware features," *J. Comput. Virol. Hacking Tech.* vol. 15, no. 4, pp. 277–305, Dec. 2019, doi: 10.1007/s11416-019-00338-7.

Vita

Eman Shaikh was born in 1998, in Dammam, Kingdom of Saudi Arabia. She received her primary and secondary education in Dammam, KSA. She received her B.Sc. degree in Computer Engineering from Prince Mohammad bin Fahd University in 2019.

In 2020, she joined the Computer Engineering master's program in the American University of Sharjah as a graduate teaching assistant. Her research interests are in cybersecurity, security analysis, and blockchain.