Improving Security for the Internet of Things

Applications of Blockchain, Machine Learning and Inter-Pulse Interval

Donpiti (Mick) Chulerttiyawong

Doctor of Philosophy

Supervisor: Professor Abbas Jamalipour

A thesis submitted in fulfilment of the
requirements for the degree of
Doctor of Philosophy

School of Electrical and Information Engineering
Faculty of Engineering
The University of Sydney
Australia

27 August 2023

Certificate of Authorship/Originality

I certify that the work in this thesis has not been previously submitted for a degree nor has it been submitted as a part of the requirements for other degree except as fully acknowledged within the text.

I also certify that this thesis/report has been written by me. Any help that I have received in my research and in the preparation of the thesis itself has been fully acknowledged. In addition, I certify that all information sources and literature used are quoted in the thesis.

[Signed]

Abstract

The Internet of Things (IoT) is a concept where physical objects of various sizes can seamlessly connect and communicate with each other without human intervention. The concept covers various applications, including healthcare, utility services, automotive/vehicular transportation, smart agriculture and smart city. The number of interconnected IoT devices has recently grown rapidly as a result of technological advancement in communications and computational systems. Consequently, this trend also highlights the need to address issues associated with IoT, the biggest risk of which is commonly known to be security. This thesis focuses on three selected security challenges from the IoT application areas of connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). For each of these challenges, a novel and innovative solution is proposed to address the identified problems. The research contributions of this thesis to the literature can be summarised as follows:

- A blockchain-based conditionally anonymised pseudonym management scheme for CAVs, supporting multi-jurisdictional road networks.
- A Sybil attack detection scheme for IoFT using machine learning carried out on intrinsically generated physical layer data of radio signals.
- A potential approach of using inter-pulse interval (IPI) biometrics for frequency hopping to mitigate jamming attacks on HBICS devices.

Details of these three studies are briefly given in the paragraphs below.

To introduce the first study, a feasible approach commonly discussed in the literature for mitigating location privacy threats for CAVs is the use of pseudonyms instead of real vehicle identifications. However, for relevant authorities to be able to identify misbehaving vehicles through their pseudonyms, it is essential that the privacy

protection mechanisms only allow for conditional anonymity and not complete anonymity. This study proposes the use of a permissioned consortium blockchain system with smart contract feature to facilitate secure and conditional privacy-preserving vehicular pseudonym issuance and management in a multi-jurisdictional road network. The use of a permissioned consortium blockchain helps mitigate security risks associated with the complexities in interorganisational data handling, such as in the areas of access control, data integrity, confidentiality, and availability. The proposed system architecture takes advantage of the predicted wide availability of Roadside Units (RSUs), and the highly viable, flexible and mature Public Key Infrastructure (PKI) technology for usage in vehicular pseudonymous communications. A small-scale simulation of the proposed architecture was successfully carried out using the Vehicles in Network Simulation (Veins) platform for integrated traffic and network simulation services (SUMO as the traffic simulator and OMNeT++ as the network simulator), and the Hyperledger Fabric platform as the permissioned consortium blockchain system. Simulation and performance analysis results reveal the feasibility of practical deployment of the scheme, and show that the scheme addresses the identified shortfalls of existing works, including the ability to achieve a better balance between connectivity and storage requirements.

The second study concentrates on the Sybil attack security threat, which refers to the situation when a malicious node falsely claims to have numerous identities. Due to the recent increase usage of unmanned aerial vehicles (UAVs) in various applications, the Sybil attack has been identified as a threat to the flying ad hoc network (FANET) paradigm and its integration with the IoT to form the IoFT. This study proposes an intelligent Sybil attack detection approach for FANETs-based IoFT using physical layer characteristics of the radio signals emitted from the UAVs as detected by two ground nodes. A supervised machine learning approach is employed and experimented with several different classifiers available in the Weka workbench platform. The experiment was carried out based on two features of the radio signals, namely, the

received signal strength difference (RSSD) and the time difference of arrival (TDoA). Simulation results revealed that the proposed scheme can achieve a high correct classification accuracy of above 91% on average, even for smart malicious nodes with power control capability operating at power levels not directly trained. In addition to the high performance, the proposed scheme is also less susceptible to various attacks commonly carried out on the upper layers, such as data spoofing, due to the use of only intrinsically generated physical layer data. Furthermore, no additional communications overheads of the UAV nodes are required for the functionality of this scheme.

Finally, the third study is relating to the security of human wearable and implantable devices in HBICS. More specifically, the use of physiological biometrics such as the timing between heartbeats, also known as the IPI, has been well-researched for mitigating threats to confidentiality and integrity; however, not quite so for the mitigation of threats to availability. The jamming of communication links to cause denial-of-service (DoS) is one such type of threat to availability. This study proposes, simulates and analyses four alternative algorithms which use IPI to add another layer of protection to the traditional pseudorandom frequency hopping system, to mitigate jamming attacks on communication links. The results reveal the feasibility for some of the algorithms to be used.

Acknowledgements

Firstly, I would like to thank my supervisor Professor Abbas Jamalipour for the tremendous support and guidance he provided throughout my PhD candidature. His support, advice and leadership are invaluable and most sincerely appreciated, especially given that I started my candidature at an unprecedented time when the world was affected by the COVID-19 pandemic.

I would like to also thank other previous and current PhD students at the Wireless Networking Group (WiNG) Lab, School of Electrical and Information Engineering, University of Sydney for their support and for passing on relevant knowledge. I would like to especially thank Dr Forough Shirin Abkenar for the warmest welcome she gave me at a time when I needed the most.

List of Publications

Journal Papers

J-1. **D. Chulerttiyawong** and A. Jamalipour, "A Blockchain Assisted Vehicular Pseudonym Issuance and Management System for Conditional Privacy Enhancement," in *IEEE Access*, vol. 9, pp. 127305-127319, 2021, doi: 10.1109/ACCESS.2021.3112013.

J-2. **D. Chulerttiyawong** and A. Jamalipour, "Sybil Attack Detection in Internet of Flying Things-IoFT: A Machine Learning Approach," in *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 12854-12866, 2023, doi: 10.1109/JIOT.2023.3257848.

Other Journal Papers

J-3. F. Shirin Abkenar, P. Ramezani, S. Iranmanesh, S. Murali, **D. Chulerttiyawong**, X. Wan, A. Jamalipour and R. Raad, "A Survey on Mobility of Edge Computing Networks in IoT: State-of-the-Art, Architectures, and Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2329-2365, 2022, doi: 10.1109/COMST.2022.3211462.

Conference Papers

C-1. **D. M. Chulerttiyawong** and A. Jamalipour, "Frequency Hopping Sequence Determination Using Inter-Pulse Interval for Human Body Interface and Control Systems," *ICC 2022 - IEEE International Conference on Communications*, Seoul, Republic of Korea, 2022, pp. 5292-5297, doi: 10.1109/ICC45855.2022.9838529.

Authorship Attribution Statement

Chapter 3 of this thesis is based on article J-1 under List of Publications. I designed the study, carried out analysis and simulation activities, and wrote the manuscript. Abbas Jamalipour was the research supervisor, while I was the corresponding author for this study.

Chapter 4 of this thesis is based on article J-2 under List of Publications. I designed the study, carried out analysis and simulation activities, and wrote the manuscript. Abbas Jamalipour was the research supervisor, while I was the corresponding author for this study.

Chapter 5 of this thesis is based on article C-1 under List of Publications. I designed the study, carried out analysis and simulation activities, and wrote the manuscript. Abbas Jamalipour was the research supervisor, while I was the corresponding author for this study.

[Signed]

Donpiti (Mick) Chulerttiyawong

Date: 27/03/2023

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

[Signed]

Professor Abbas Jamalipour

Date: 31/03/2023

Contents

List of Figures

Abbreviations

ABE	Attribute-Based Encryption
AoA	Angle of Arrival
BLE	Bluetooth Low Energy
CA	Certification Authority
CAN	Controller Area Network
CAV	Connected and Autonomous Vehicle
CFT	Crash Fault Tolerant
CNN	Convolutional Neural Network
CRL	Certificate Revocation List
CSR	Certificate Signing Request
DAG	Directed Acyclic Graphs
DDoS	Distributed Denial-of-Service
DL	Deep Learning
DoS	Denial-of-Service
DRL	Deep Reinforcement Learning
ECC	Elliptical Curve Cryptography
ECDSA	Elliptic Curve Digital Signature Algorithm
ECG	Electrocardiogram
FANET	Flying Ad Hoc Network
GAN	Generative Adversarial Network
GCS	Ground Control Station
GPS	Global Positioning System
HBICS	Human Body Interface and Control Systems
ID-PKC	Identity-Based Public Key Cryptography
IDS	Intrusion Detection Systems

IoFT	Internet of Flying Things
IoT	Internet of Things
IoV	Internet of Vehicles
IPI	Inter-Pulse Interval
IPS	Intrusion Prevention Systems
ITS	Intelligent Transportation Systems
LIN	Local Interconnect Network
MANET	Mobile Ad Hoc Network
MITM	Man-in-the-Middle
ML	Machine Learning
mmWave	Millimeter-Wave
MOST	Media Oriented Systems Transport
OBD	On-Board Diagnostics
OBUs	On-Board Unit
PBFT	Practical Byzantine Fault Tolerance
PKI	Public Key Infrastructure
PoS	Proof-of-Stake
PoW	Proof-of-Work
PPG	Photoplethysmogram
PUF	Physical Unclonable Function
RAM	Random Access Memory
RKES	Remote Keyless Entry Systems
RL	Reinforcement Learning
RNN	Recurrent Neural Network
RSS	Received Signal Strength
RSSD	Received Signal Strength Difference
RSU	Roadside Unit
Rx	Receive

TDoA	Time Difference of Arrival
ToA	Time of Arrival
Tx	Transmit
UAS	Unmanned Aircraft Systems
UAV	Unmanned Aerial Vehicle
V2G	Vehicle to Grid
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
V2X	Vehicle to Everything
VANET	Vehicular Ad Hoc Network
VLC	Visible Light Communication
WBAN	Wireless Body Area Networks

Chapter 1 Introduction

This thesis focuses on improving security for the Internet of Things (IoT), which has been and is still a very active research area. This chapter introduces the IoT concept and the associated security challenges, with emphasis on three application domains that will be explored in subsequent chapters, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). Furthermore, this chapter also outlines the motivations and objectives, as well as the contributions of the research works carried out in later chapters of this thesis.

1.1. Internet of Things

The Internet of Things (IoT) is a concept originally introduced in 1999 by Kevin Ashton, describing physical objects of various sizes that can seamlessly connect and communicate with each other without human intervention [1]. Since the concept's introduction, the amount of interconnected devices has grown rapidly due to technological advancement in communications and computational systems [2]. The types variety of interconnected IoT devices has also expanded to cover various application areas, including healthcare, utility services, automotive/vehicular transportation, smart agriculture and smart city [3]. Consequently, IoT devices expanded to include things like blood pressure meters, home televisions, restaurant refrigerators, connected cars, robotic manufacturing systems, smart grid, and countless industrial control systems [4]. According to the latest global forecast and analysis reporting by Cisco, it is estimated that by 2023, there will be 29.3 billion networked devices, which is equivalent to approximately 3.6 devices per person when taking into account the world population. The same report also states that the number

of networked devices in 2018 was 18.4 billion devices, or 2.4 devices per person, indicating significant growth within just half a decade [5].

As can be seen, IoT is a broad concept covering a variety of different application domains. As depicted in Figure 1.1, the focus of this thesis will be on three application domains, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). Brief introductions of these application domains are given in the following subsections.

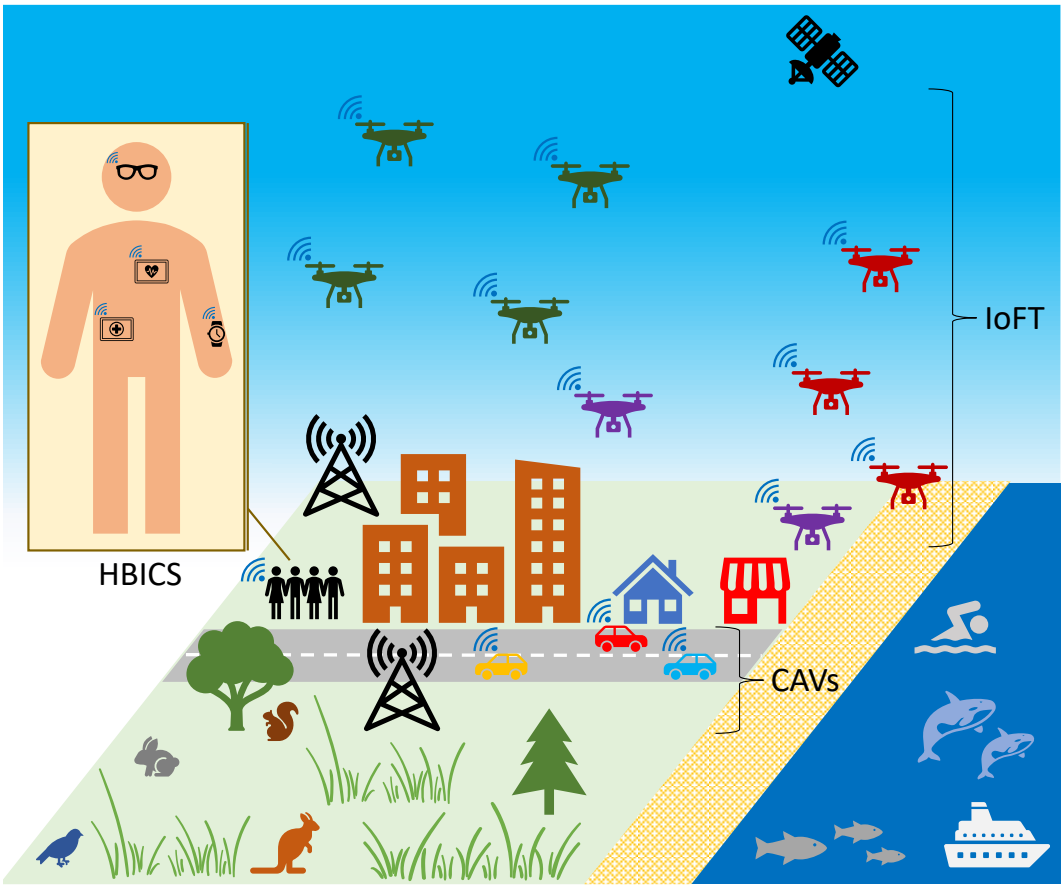

Figure 1.1: IoT Application Domains Focused in This Thesis

1.1.1. Connected and Autonomous Vehicles

The connected and autonomous vehicles (CAVs) concept is known to be forming the backbone of future next-generation intelligent transportation systems (ITS) to provide travel comfort, road safety, and other value-adding services [6]. Due to CAVs requiring to be equipped with many sensors, there are also additional requirements for these sensors to communicate critical sensory information with other systems, both internally and externally, such as to other nearby vehicles. Consequently, there appears to be more and more research on vehicular communications concurrently with the evolution of various related technologies.

Apart from the term “connected and autonomous vehicles (CAVs)”, there are also many different descriptive terms used in the literature that are associated with vehicular communications, such as Vehicular Networks, Vehicular Ad Hoc Networks (VANETs), Internet of Vehicles (IoV), Connected Vehicles, Vehicle to Vehicle (V2V), Vehicle to Infrastructure (V2I), Vehicle to Grid (V2G) and Vehicle to Everything (V2X) [6] [7] [8]. Although these terms may have some slight differences in meaning, they share a common element, which is that communications systems are required between entities, whether it be internal or external to the vehicle, to connect them to each other.

Prominent radio access technologies for vehicular communications include DSRC and cellular networks (LTE-V/4G/5G) [7] [9] [10]. It is envisioned that connected cars to be released to the world’s public roadways will first be using DSRC technology [11]. This is because DSRC is a mature technology that has been developed for over ten years [9]; however, there is also a lack of clear plan for future evolution of the standards. Conversely, LTE-V appears to have a brighter future going forward, as although it is still in the early stage of development, it has several advantages such as more bandwidth, wider coverage and the ability to reuse existing cellular network

infrastructure. Emerging radio access technologies such as millimeter-wave and visible light are also discussed in the literature as potential new trends for vehicular communications [10].

Vehicular communications has some special characteristics that are unique among other IoT systems. As the vehicle's communications systems are assumed to be powered by the more abundance drivetrain power, there would be less constraints on power availability; in fact, several studies such as Obaidat et al. [12] and Arif et al. [13] infer that this means there is no power constraint. There is also a higher degree of mobility pattern predictability. Vehicular networks can also potentially be highly dynamic because of high mobility and the number of vehicles joining and leaving. Due to the characteristics of vehicular networks, it is not surprising to see that many research challenges in vehicular communications are associated with mobility. Furthermore, the expected high number of vehicles means that the network is predicted to be quite large in scale [7].

1.1.2. Internet of Flying Things

The Internet of Flying Things (IoFT) concept is a relatively new research area that is built upon the integration of Unmanned Aerial Vehicles (UAVs) and IoT [14] [15]. UAVs are pilotless aerial vehicles either autonomously controlled by a computer or remotely controlled by a pilot on the ground. They are also commonly known as drones, and have recently gained increasing usage in civilian applications, including agricultural remote sensing, search and rescue operations, disaster monitoring, weather monitoring, pollutant studies and delivery of products such as food. These relatively new applications are in addition to their military applications, such as strike, reconnaissance and border surveillance, which have been in place since several decades back. With such increasing usage, the global UAV market compound growth rate is estimated to be as strong as 19.99% between 2016 to 2022 [16].

UAV is known to be an important element of the Unmanned Aircraft Systems (UAS), together with the other two elements being the ground control station (GCS) and the communications links [14] [15]. To enable more effective operations, different UAVs are often operated together in different arrangements. For example, they can be coupled together using physical links. Alternatively, they can be arranged in a formation where the UAVs are not physically connected but their relative motions are constrained to keep such a formation. Another example is homogenous teams of UAVs operating in swarms [16]. The UAV swarm arrangement has especially been increasingly discussed as an operation model of great potential for various applications, including search and rescue operations [17] [18] and air quality index monitoring [19].

The UAVs communications network is known as a flying ad hoc network (FANET), which is a subclass of mobile ad hoc networks (MANETs), but with nodes possessing aviation characteristics [14]. In the FANET paradigm, UAVs can communicate with each other without requiring an access point, given that at least one of them connects to a GCS or a satellite [16]. This allows the UAV nodes to cooperate through ad-hoc networking, enabling the achievement of operations requiring higher scalability, reliability, survivability, and a lower cost [14].

Given the aviation nature of FANET, there exist several unique characteristics. Firstly, the nodes are expected to be of relatively high mobility. Multiple connections among these nodes are also anticipated. Consequently, from the perspective of the network topology, frequent change is also an expectation. Nevertheless, node density is projected to be much lower than some other paradigms, such as VANET. In terms of radio propagation, line-of-sight (LoS) links between the nodes and the GCS can generally be assumed. As for the power consumption, this is limited by the energy source of the specific UAV, where battery capacity can vary subjecting to the UAV type and size [14] [15] [16]. Due to the uniqueness of FANET, several challenges have been

identified, including in the areas of networking protocols, mobility models, security, quality of service and standards [16].

From the perspectives of radio frequency and access technology, FANET uses a variety of these depending on applications. The frequencies can range from lower ones such as below 1 GHz, which is suitable for control links and telemetry, to higher ones such as above 5 GHz, where quality video broadcasts can be carried out. Due to the different FANET applications and communications ranges, various radio access technologies are expected to be used, such as multi-hop IEEE 802.11, Zigbee, WiMAX, EDGE, UMTS and LTE [16].

1.1.3. Human Body Interface and Control Systems

The concept of human body interface and control systems (HBICS) refers to the information exchange between devices inside, on, and within the proximity of a human body. It is also known as wireless body area networks (WBAN). As the names suggest, the focus of the concept is substantially on human wearable and implantable devices. Due to such nature of the concept, HBICS cover both medical and non-medical applications. Medical applications can include the monitoring and control of health conditions, such as fatigue, asthma, diabetes, cardiovascular diseases and cancer detection. As for non-medical applications, these can range from things like entertainment to non-medical emergency management and security management [20].

Currently, HBICS implantable devices appear to be primarily for medical applications. However, wearable devices tend to cover a broader range of applications. For example, Seneviratne et al. [21] classified wearable devices into three categories, namely, accessories, e-textile and e-patch. Accessories can be broken down further into wrist-worn devices, head-mounted devices and other accessories. Wrist-worn devices include things like smart watches and wristbands. Examples of head-mounted

devices include smart eyewear, headsets and earbuds. Examples of other accessories include smart jewellery and various straps for health tracking and other functionalities. As for e-textile, items that fall into this category include smart clothing garments such as shirts, pants and undergarments, as well as foot and hand-worn items like shoes, socks and gloves. Finally, the e-patch category includes items such as sensor patches and e-tattoo/e-skin.

Similar to many other IoT systems, challenges associated with HBICS devices stem from the limitations in resources due to the constraints associated with the intended operation scenarios. These resources are such as energy, storage, computing and communication [22] [23]. Nevertheless, safety and security are often cited as concerns that are especially important to HBICS when compared to some other IoT application domains, due to the possibility of direct involvement in life-critical information and potentially in hostile environments [20] [23] [24] [25].

Communications in HBICS can generally be classified into three to four tiers [24] [26] [27]. Tier-1 generally refers to communications where both the sender and the receiver are inside a human body. Tier-2, Tier-3 and Tier-4 then gradually refer to communications from a human body to an off-body device and beyond. In terms of communications protocols, HBICS operational scenarios generally favour those that support low power, such as Bluetooth Low Energy (BLE), IEEE 802.15.4, IEEE 802.15.6 and IEEE 802.11ah [21] [24] [28].

1.2. Security for the Internet of Things

With the trend of increasingly interconnected IoT devices also comes the exacerbation of issues associated with IoT, the biggest risk of which is commonly known to be security issues [29]. Mitigating IoT security issues is quite challenging as most IoT devices are designed to be small in size and have inherently limited resources (i.e.,

battery, processing, and storage) [1]. Research in IoT security is also known to still be in the conceptual stage, requiring further exploration to develop new innovative security solutions [29].

Another challenge for the security of IoT is the associated heterogeneous characteristic. There exists a large variety of IoT technologies, including protocols and standards [30] [31]. This leads to the lack of interoperability between systems, a reduced number of skilled personnel available to contribute to works in each technology type, and a reduced number of solutions with cross-domain applicability [30].

As outlined in Section 1.1, this thesis focuses on three IoT application domains, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS), each of which is a broad field of study in its own right. As a consequence, each application domain has associated security challenges that are more or less unique, but often still share some similar characteristics with those found in other domains. This thesis explores three selected security challenges, each of which belongs to an application domain mentioned above. It is the author's aspiration that the studies carried out and captured in this thesis will contribute to enabling more secure IoT devices integration, and subsequently lead to more effective and robust use of IoT technology.

1.3. Research Motivations and Objectives

As outlined in Section 1.1, the number of interconnected IoT devices has recently grown rapidly due to technological advancement in communications and computational systems. Consequently, IoT devices are now quite ubiquitous, as well as becoming an integral part of human lives. Since interconnected IoT devices are used in many different application areas, this phenomenon brings about tremendous

benefits to humanity. Nevertheless, with such trends also come the associated challenges awaiting to be tackled to safeguard the benefits realisation, one interesting area of which is security. Due to the rapid growth of the IoT paradigm, there are still a lot of open challenges as to how to improve IoT security. These elements collectively form the motivation factors for the studies carried out in this thesis, with the primary objective being that the solutions developed hereby would lead to a world where the IoT system is robust, effective and safe to use, which would then lead to the enhancement of various other societal developments.

With IoT being such an enormous area of study, this thesis cannot address the improvement of security in all domains. Notwithstanding that, the literature review on IoT security carried out in Chapter 2 has identified three interesting challenges, which are individually addressed in Chapter 3, Chapter 4 and Chapter 5 of this thesis. To categorise, the selected three challenges are in the areas of connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). Additionally, the literature review also identified ideas, techniques, and most notably technologies such as blockchain and machine learning systems, which may be useful in contributing to the formation of potential solutions. The motivations and objectives of the three studies captured in this thesis are briefly given below.

Firstly, a challenge identified relating to vehicular communications is location privacy preservation. Although the use of pseudonyms instead of real vehicle identifications has been known as a mitigation for this problem, there is a trade-off for this with non-repudiation. To resolve such a trade-off, the use of vehicular pseudonyms needs to be made conditional, allowing the real identity to be traceable by authorised parties. In this context, the distinct immutable distributed ledger property of blockchain systems is a potential solution to assist in the management of conditionally anonymised vehicular pseudonyms. This is especially so for the consortium blockchain type, which has great potential in the provision of a secured integrated solution for a road network

jointly managed by different jurisdictions, such as different states within a country. The identified challenge and the existence of consortium blockchain systems as a potential solution enabling technology, thus, motivate the development of a novel pseudonym issuance and management system.

The next challenge identified is related to IoFT security. More specifically, the recent trend of increasing UAV usage in civilian applications, in addition to the traditional deployment in the military domain since several decades back, has highlighted the need for additional countermeasures against security threats, including the Sybil attack. In this context, the use of machine learning has been identified as a tool of high potential to intelligently detect Sybil attack instances in FANETs-based IoFT. Consequently, this challenge, and the existence of machine learning systems as a potential solution enabling technology, stimulate the invention of a state-of-the-art Sybil attack detection scheme.

Finally, the last challenge identified is related to HBICS security. The literature review conducted indicates that numerous authentication and encryption applications that use biometrics, such as inter-pulse interval (IPI), have been developed to address various security challenges in the past. These developments were largely in conjunction with the fuzzy commitment scheme, which allows for errors in what is equivalent to a decryption key, to be tolerable to a certain degree. On the other hand, denial-of-service (DoS) is known to be a major attack type in HBICS, and one method of carrying out such an attack is through the launch of wireless communications link jamming. In this context, the existing solutions, which would have worked fine with authentication and encryption applications, cannot be adopted for use in frequency hopping applications due to the fundamental difference in how frequency hopping operates. This gap triggers the development of a new frequency hopping approach that uses IPI biometrics to add another layer of protection to the traditional pseudorandom frequency hopping system, potentially bringing enormous benefits to HBICS.

1.4. Thesis Contributions

The contributions of this thesis can be divided into three areas. Firstly, the contributions associated with the proposed pseudonym issuance and management system outlined in Chapter 3. Secondly, the contributions resulting from the development of the novel IoFT Sybil attack detection scheme proposed in Chapter 4. Thirdly, the contributions arose from the development of potential solutions to enable IPI biometrics to be used for frequency hopping sequence determination, as outlined in Chapter 5.

The main contributions of Chapter 3 are summarised as follows:

- To fill a knowledge gap in the literature relating to conditionally anonymised vehicular pseudonym management in a multi-jurisdictional road network.
- To achieve conditionally anonymised vehicular pseudonym management supporting multi-jurisdictional road networks, while also concurrently:
    - minimising the associated complexities and security risks at interfaces between different jurisdictions; and
    - enabling integrated collaboration between different organisations.
- To investigate and demonstrate the use of permissioned consortium blockchain paired with the traditional PKI-based cryptography system in carrying out pseudonym issuance and management in a dynamic, secure, conditional privacy-preserving and distributed manner.

The main contributions of Chapter 4 are summarised as follows:

- To fill a knowledge gap in the literature relating to Sybil attack detection in FANETs-based IoFT which is still quite deficient in general.
- To achieve Sybil attack detection in FANETs-based IoFT using intrinsically generated physical layer data of radio signals emitted from the UAVs.

Advantages associated with this are such as less susceptibility to attacks involving information spoofing and not requiring additional communications overheads.

- To achieve Sybil attack detection in FANETs-based IoFT, where both classic malicious nodes with fixed power and smart malicious nodes with power control capability may be presented.
- To investigate and demonstrate the use of machine learning in carrying out Sybil attack classification determination based on two attributes, namely RSSD and TDoA ratios of two different radio signals, obtained using only two monitoring nodes.

The main contributions of Chapter 5 are summarised as follows:

- To fill a knowledge gap in the literature relating to the use of IPI biometrics as an added layer of protection to the traditional pseudorandomly determined frequency hopping pattern.
- To investigate and demonstrate the derivation of IPI biometrics frequency hopping determinants that are tolerant to noise and errors caused through IPI measurements being taken at different parts of the body.

1.5. Thesis Organisation

This chapter outlines an overview of the Internet of Things (IoT) concept, focusing on three application domains, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). It also briefly introduces the topic of IoT security, laying the foundation for further discussions in Chapter 2 and subsequent chapters. Furthermore, research motivations and objectives, as well as a summary of the contributions of this thesis

are also discussed in this chapter. As for the remaining chapters, they are organised as follows.

In Chapter 2, a comprehensive literature review on the topic of IoT security is given. More specifically, in Section 2.1, general IoT security threats are discussed. This follows by Section 2.2 which discusses IoT security by application domain, including vehicular communications security, Internet of Flying Things (IoFT) and human body interface and control systems (HBICS). Section 2.3 then discusses IoT security solution enabling technologies, including blockchain and machine learning. Finally, Section 2.4 outlines a brief introduction to quantum technology and how its materialisation is likely to affect the future of IoT quite significantly.

Chapter 3 proposes the use of a permissioned consortium blockchain system with smart contract feature to facilitate secure and conditional privacy-preserving vehicular pseudonym issuance and management in a multi-jurisdictional road network. More specifically, Section 3.1 outlines an introduction to connected and autonomous vehicles and the associated vehicular location privacy threats, especially in the context of a multi-jurisdictional road network. In Section 3.2, the existing related works are reviewed, including in the areas of vehicular pseudonyms, roadside units (RSU) and blockchain technology. Section 3.3 then outlines the motivations and contributions of the chapter. This is followed by Section 3.4, which outlines the proposed vehicular pseudonym management system architecture. Section 3.5 outlines the simulation environment, including the use of the Veins vehicular network simulation framework and the Hyperledger Fabric permissioned consortium blockchain platform. Subsequently, the simulation results are discussed and evaluated in Section 3.6. Finally, the chapter concludes in Section 3.7.

Chapter 4 outlines a state-of-the-art intelligent Sybil attack detection approach for FANETs-based IoFT using supervised machine learning carried out on two physical layer features of the radio signals emitted from UAV nodes, namely, the received

signal strength difference (RSSD) and the time difference of arrival (TDoA). More specifically, Section 4.1 outlines an introduction to the increasing usage of UAVs and how this trend accentuates various security threats, including the Sybil attack. In Section 4.2, the existing related works are reviewed, including on the use of physical layer features of the radio signal for positioning systems, the existing Sybil attack detection approaches and the use of machine learning for IoT security. Section 4.3 then outlines the motivations and contributions of the chapter. This is followed by Section 4.4, which outlines the proposed intelligent Sybil attack detection scheme. Section 4.5 outlines the simulation environment, including the OMNeT++/INET simulator and the Weka machine learning workbench platform. Subsequently, the simulation results are discussed and evaluated in Section 4.6. Finally, the chapter concludes in Section 4.7.

Chapter 5 presents a novel frequency hopping approach in HBICS, which uses inter-pulse Interval (IPI) biometrics to add another layer of protection to the traditional pseudorandom frequency hopping system. More specifically, Section 5.1 outlines an introduction to HBICS and the associated security matters, including how biometrics are known to be potentially suitable for various security applications. In Section 5.2, the existing related works are reviewed, including the use of IPI for security applications and the use of frequency hopping to counteract frequency jamming attacks. Section 5.3 then outlines the motivations and contributions of the chapter. This is followed by Section 5.4, which outlines the proposed algorithms to be used with IPI biometrics data. Section 5.5 outlines the simulation environment, including the source of IPI data and the MATLAB/Simulink simulation platform. Subsequently, the simulation results are discussed and evaluated in Section 5.6. Finally, the chapter concludes in Section 5.7.

Chapter 6 is the concluding chapter of this thesis. It summarises all the studies carried out and the contributions they make to the literature. Remaining challenges and future research directions are also briefly outlined.

Chapter 2 Internet of Things Security: A Literature Review

Before proceeding to derive new solutions for Internet of Things (IoT) security issues, it needs to be known what problems exist and what solutions have already been developed. These findings will then identify gaps where the creation of novel security solutions would be beneficial. To progress forward, the literature needs to be thoroughly reviewed, as summarised in this chapter. In this context, references to relevant articles are also provided throughout the chapter, and interested readers are strongly encouraged to refer to these for more comprehensive reading. As illustrated in Figure 2.1, the literature review covers both the topic of general IoT security threats, as well as more specific emphasis on IoT security by application domain. Additionally, it also encompasses the topics of blockchain and machine learning from the aspect of their use as IoT enabling technologies. Furthermore, it also briefly covers the area of quantum security, which will likely play an important role in the future of IoT.

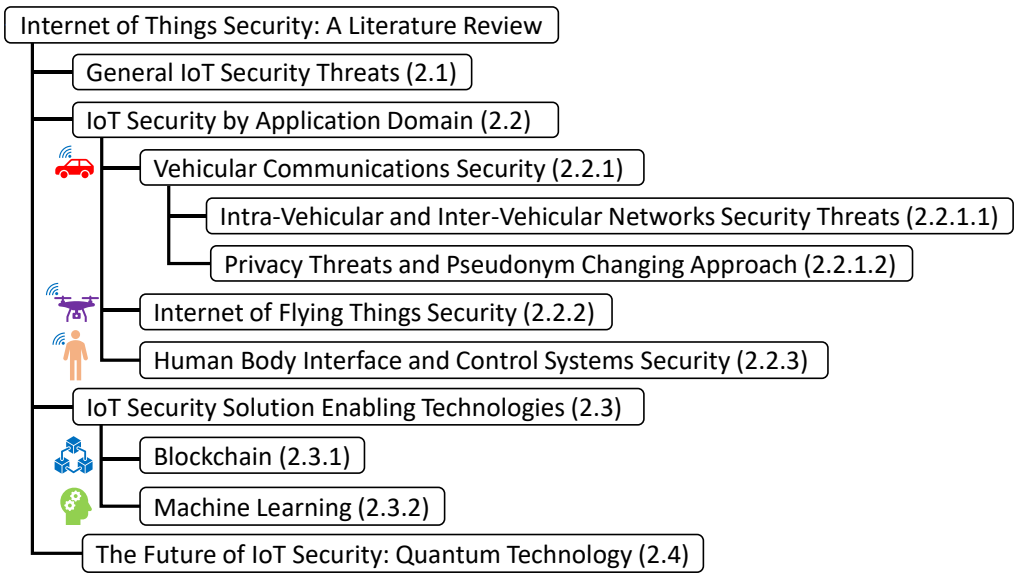

Figure 2.1: Literature Review Organisation

2.1. General IoT Security Threats

Common IoT security requirements in the literature include attributes such as confidentiality, integrity, authentication, accountability, availability, non-repudiation and privacy [3] [32] [33]. Privacy is an interesting concept, as it is frequently discussed as a separate entity, but is still often used alongside the concept of security. For example, Arif et al. [13] and van Der Heijden et al. [34] discussed the concept of trading off between security and privacy. However, as inferred by authors such as Eckhoff and Wagner [35], there is no privacy in the absence of security. Therefore, it is most logical to consider privacy as being one of the security attributes. When the trade-off between security and privacy is discussed, it can be viewed as a trade-off between privacy and other security attributes.

Many research papers on IoT security have divided IoT security architecture into three layers: *perception*, *transportation*/*network*, and *application*, such as defined in Zhao and Ge [36] and Jing et al. [37]. A review of recent survey papers suggests that there are currently various challenges associated with all layers, including those that may potentially be beneficial if solved using cross-layered solutions.

Security attacks in IoT have been classified into different types by various authors. Common types of attacks in IoT are summarised in Table 2.1, following a review of various publications, including Ramezan et al. [33], Butun et al. [38] and Chaabouni et al. [31].

Table 2.1: Common IoT Security Attack Types Summarised From Ramezan et al. [33], Butun et al. [38] and Chaabouni et al. [31]

Attack Type	**Description**
Denial-of-Service (DoS) / Distributed Denial-of-Service (DDoS)	Causing losses in the availability of services. This may be achieved through several means, such as: • Jamming (e.g., transmit on the same radio frequency) • Flooding (e.g., sending excessive data to a node)

Attack Type	Description
	• Energy depletion (e.g., keeps sending unnecessary information to a battery powered node to prevent it from going to sleep mode) DDoS refers to when DoS is carried out through multiple compromised systems.
Eavesdropping/Traffic Analysis	Passive attack that listens to and gathers information transmitted over a communication channel for analysis.
Hole Attacks (e.g., blackhole, greyhole, sinkhole and wormhole)	When a node manipulates the received messages in one or more ways that deviate from the behaviours expected by other nodes: • Blackhole: Node drops the received messages. • Greyhole: Node selectively drops some received messages. Because only some messages are affected, this can be difficult to detect. • Sinkhole: Node attracts high traffic by pretending to be the optimal path for all messages to reach their destinations. • Wormhole: Two nodes working together to form a secret tunnel and advertised as neighbours even though they may actually be far apart.
Man-in-the-Middle (MITM)	Intercepting communication channel to eavesdrop (e.g., for the purpose of launching a replay attack later).
Replay	Attacker eavesdrops information and retransmits this at a later time (e.g., retransmits a user's credentials to gain unauthorised access to a system).
Spoofing/Impersonation	Attacker generates illegitimate imitated information.
Sybil	A single node pretends to be multiple identities, causing confusion to other nodes (e.g., in an attempt to gain an unfair advantage in the voting / trust evaluation process).

According to Nespoli et al. [39], there are four phases in the cyber defence cycle, where each feeds the next phase and forms a loop back to the start, namely and in the order of: *prevention*, *detection*, *reaction* and *forensics*. There are numerous potential mechanisms in the literature to defend against security attacks in IoT. For example, Intrusion Prevention Systems (IPS) can be used for the prevention phase, Intrusion Detection Systems (IDS) can be used for the detection phase, and the robust design of system architectures and protocols can be used to mitigate threats in all phases. Auditing and logging activities are also collectively known as one of the

traditional defence mechanisms [31], and can be used to support activities carried out in most, if not all, phases. Nevertheless, due to the previously discussed IoT characteristics, such as the continuing technological advancement, the inherently limited resources available, and the diversity in protocols and standards, there is still much room for improvement in IoT security.

2.2. IoT Security by Application Domain

This section captures the literature review of some IoT security issues categorised by application domains being focused on in this thesis. Firstly, the topic of vehicular communications security is reviewed, in preparation for the content of Chapter 3. Next, the topic of the Internet of Flying Things (IoFT) security is explored, in preparation for the content of Chapter 4. Finally, security threats to human body interface and control systems are given, in preparation for the content of Chapter 5.

2.2.1. Vehicular Communications Security

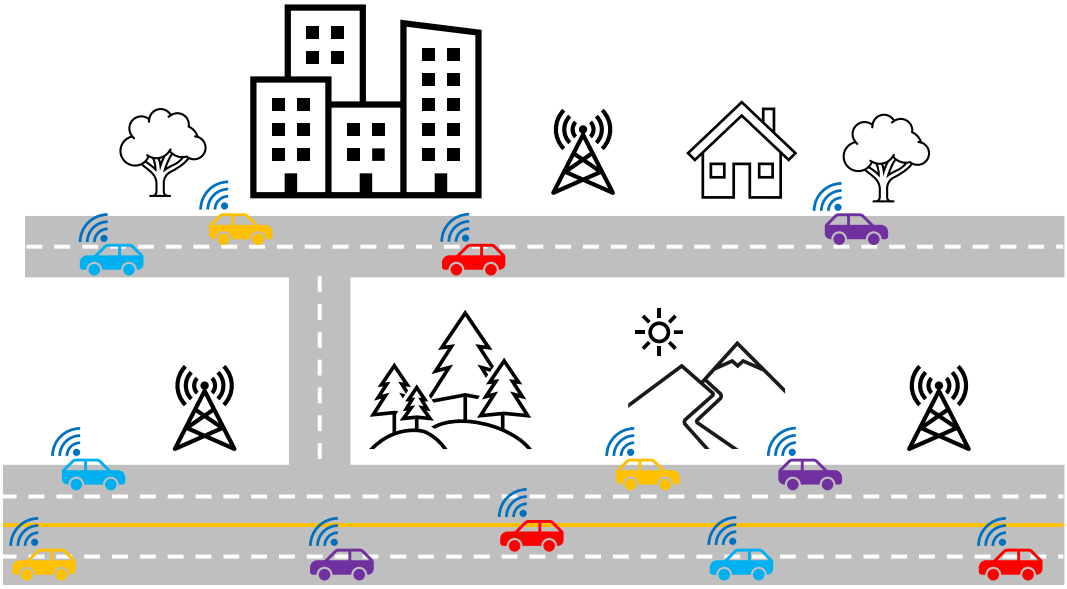

Figure 2.2: Road Network With Connected Vehicles and Infrastructures

Consider a road network with connected vehicles and infrastructures as shown in Figure 2.2. There are numerous potential security issues in this scenario, both those that are applicable to communications internally within the vehicle and externally. Security, including privacy, is known to be a challenge in vehicular communications [40] [41]. Contributory factors include the heightened need for protection, as well as the lack of standardised security guidelines, architectures and protocols [7] [41]. In addition, security is considered to be an afterthought in many systems [42]. Therefore, it is imperative that security gets considered early in the design of vehicular communications systems. The following subsection explores some challenging security issues commonly discussed in the literature, including both intra-vehicular and inter-vehicular threats. The subsection after then explores more specifically into the topic of location privacy threats and the use of pseudonym changing approach as a mitigation technique.

2.2.1.1. Intra-Vehicular and Inter-Vehicular Networks Security Threats

As most vehicles internal systems are connected using bus network topology [7], intra-vehicular networks security threats usually involve bus system exploitation. Common bus systems include Controller Area Network (CAN), Local Interconnect Network (LIN), FlexRay, Media Oriented Systems Transport (MOST) and Ethernet. Interfaces that enable attacks include physical ports, such as On-Board Diagnostics (OBD) ports, USB ports, and ports attached to electric vehicle charging systems, as well as infotainment and telematics systems [43]. Potential attacks are such as impersonation, eavesdropping, injection, replay, denial-of-service (DoS), and bus-off attacks [43] [44].

Inter-vehicular communication refers to communication between the vehicle On-Board Units (OBUs) and other external entities, including other vehicles and infrastructures, such as roadside Units (RSUs). Several terms, such as V2V, V2I and V2X, are commonly found in the literature (e.g., throughout Mahmood et al. [7] and El-Rewini et al. [43]) to describe this type of communication.

Unlike intra-vehicular networks, inter-vehicular networks security threats generally involve more exploitation of wireless communications technologies. These technologies range from Remote Keyless Entry Systems (RKES) to a wide variety of wireless communications access technologies (e.g., DSRC, Cellular, ZigBee, Bluetooth and Wi-Fi). Potential attacks include eavesdropping, Global Positioning System (GPS) spoofing, modification/replay and jamming/denial-of-service [43] [44].

2.2.1.2. Privacy Threats and Pseudonym Changing Approach

One of the major vehicular communications security concerns that get frequently discussed in the literature is the topic of threats to privacy [7] [8] [13] [34] [42] [43]

[45] [46] [47] [48] [49]. This is driven by both users preference to potentially be anonymous, as well as privacy related legislation that have been in place, or may be in place in the future, to protect users privacy. As outlined previously, mobility contributes to many vehicular communications security challenges, and privacy is one of them. Although user privacy issues may be common in other types of IoT systems, location privacy is an additional aspect that is more applicable to vehicular communications due to the associated mobility. Preserving location privacy is known to be more challenging than that of user privacy [42].

Other related issues that need to also be considered include trade-offs between privacy and other security and functionality attributes [13] [45] [47]. For example, in order to obtain location proof, there would be an inherent trade-off between this and location privacy. There would also be potential trade-offs between privacy and liability/non-repudiation.

There are several requirements for location privacy in vehicular communications [49]. Firstly, the amount of information to be revealed by a user should just be the minimal required for ensuring the functionalities of the vehicular network. Secondly, the messages sent out should have anonymity. However, note that this is where the trade-off with liability/non-repudiation occurs. To solve this issue, anonymity should be made conditional, with the real identity still being traceable by authorised parties. Thirdly, messages from the same vehicle should not be linkable for an extended period of time. Finally, there should be perfect forward privacy, meaning that any credential revocation of a vehicle should not affect the unlinkability of any of its other credentials.

Several privacy protection mechanisms against vehicular tracking have been discussed in the literature. For more details, interested readers may refer to studies such as van Der Heijden et al. [34], Wang et al. [45], Chen et al. [46], Sharma and Kaushik [47], Manivannan et al. [48], Boualouache et al. [49], Ali et al. [50] and Khelifi et al. [41].

Among these, the most prominent approach appears to be where pseudonyms are used instead of real identifications. Nevertheless, there are also other schemes in discussion, such as those involving the use of group signatures, identity-based public key cryptography (ID-PKC), and hybrid approaches in which different schemes are combined [49] [50].

As outlined above, the pseudonym changing approach has been discussed in numerous research works. The approach involves the use of pseudonyms instead of real vehicle identifications [48]. The approach works on the principle that if vehicles attach long-term identities to outgoing messages, they can potentially be tracked using a trace of received messages and the associated location information [34]. Therefore, the use of short-term pseudonyms for the vehicle's identity has been proposed to be used to prevent tracking.

Nevertheless, simply changing pseudonyms can be ineffective and inefficient [49]. For example, such schemes may not properly address linkability issues. According to van der Heijden et al. [34], linkability is classified into four categories. Firstly, *full linkability* is where there is no pseudonym change at all, and thus all messages transmitted from the same OBU can be linked. *Explicit linkability* is where some sort of pseudonym scheme gets deployed; however, it still allows for direct access to an identity. *Implicit linkability* or *inference* is where pseudonyms can still be linked to reveal partial or complete identities, even without direct identification. Sources of information available to be used for this type of linkability include certificates that may contain attributes of a vehicle, such as length, height and colour, the message content itself, and the transmission signal properties. The final category is where there is *no linkability*. This is the ideal scenario and occurs when it is not possible to determine whether two messages originate from the same or different vehicles; however, it is nearly impossible to provide any system functionality in this scenario. Furthermore, even in the no linkability scenario, tracking is still possible by simply following a vehicle.

As can be seen from the above, schemes that use pseudonym changing approach need to be well-designed and take different types of linkability into consideration. Consequently, the use of pseudonym changing approach poses several challenges. To prevent linkability between messages from the same OBU, the time and location of when and where vehicles change their pseudonyms are important. The concepts of *mix-zone* and *radio silence periods* have been discussed in various studies (e.g., Sharma and Kaushik [47], Boualouache et al. [49] and Ali et al. [50]) to resolve these issues. The idea is that vehicles that participate in a particular pseudonym changing scheme would change their pseudonyms in the mix-zone, and after the radio silence period, to prevent linkability. However, there are challenges associated with these concepts. For example, the mix-zone may be functional on multi-lane roads but not so much on one-way roads [47], and the radio silence period may impact safety as no safety messages get broadcast during such periods [49].

Another challenge is if vehicles are required to participate in voting or other trust-based evaluations of data, the use of pseudonyms may hinder the intended operation. Therefore, a voting scheme needs to be designed to allow participants in pseudonym changing schemes to still be able to successfully participate in [34].

2.2.2. Internet of Flying Things Security

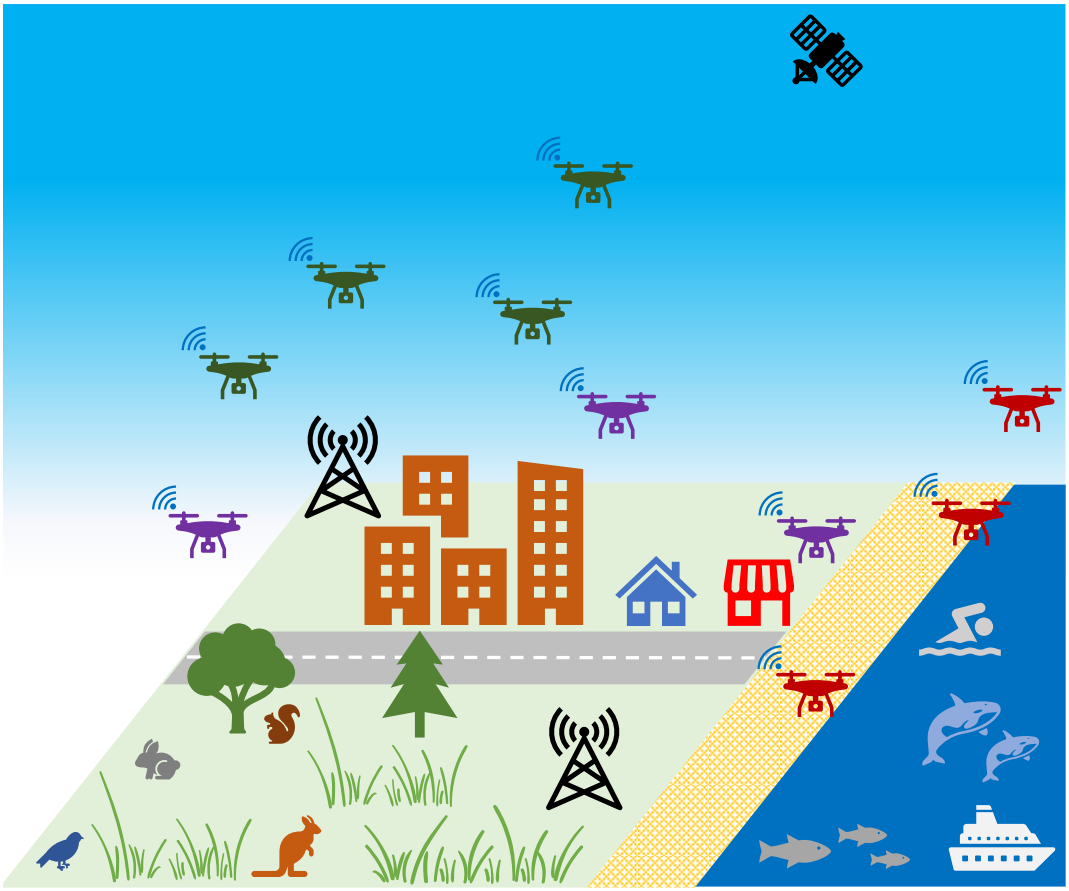

Figure 2.3: IoFT Formation Through Ubiquitous UAV Deployments

As already outlined in Section 1.1.2, the IoFT scenario, which is hereby illustrated in Figure 2.3, is becoming more prevalent due to the increasing usage of UAVs in civilian applications. There are several challenges in IoFT that are quite unique because of the traits associated with FANET, one of which is security. To elaborate, the specific nature of FANET networks is of collaborative characteristics; consequently, many wireless communications links would be expected between highly mobile UAV nodes and with GCS or satellite. The nature of FANET operations also means that the UAV nodes are not anticipated to always be connected directly to the GCS and satellite. In contrast, UAV nodes are still expected to communicate with other UAV nodes, even without access to the GCS and satellite. Such unique characteristics create an environment

with a tendency for things to become uncontrollable relatively easily, thus making securing IoFT quite challenging [14] [15] [16].

In terms of areas of focus, Chriki et al. [16] suggested consideration of seven main security criteria: availability, integrity, confidentiality, authenticity, non-repudiation, authorisation and anonymity. The authors also suggested that the main security services that an attacker would likely want to break are authentication, availability, confidentiality and integrity. Additionally, Zaidi et al. [14] recommended enhancing IoFT security and privacy in three layers, namely, application, transportation and physical layers. Furthermore, Pigatto et al. [15] highlighted the security issues associated with big data that would be brought in with IoFT, especially when such big data are expected to most likely be nonstructured.

An example of a potential security threat to IoFT is the Sybil attack. As already briefly described in Section 2.1, the Sybil attack is a well-known security threat to the IoT. More specifically, it is an authenticity threat, where a node tries to cause confusion to other nodes by pretending to be of multiple identities. In the context of FANETs, the Sybil attack is identified as a threat with several incentives, such as allowing a malicious node to illegitimately acquire more weight in a voting system and creating an illusion of traffic congestion in a particular area [51] [52].

2.2.3. Human Body Interface and Control Systems Security

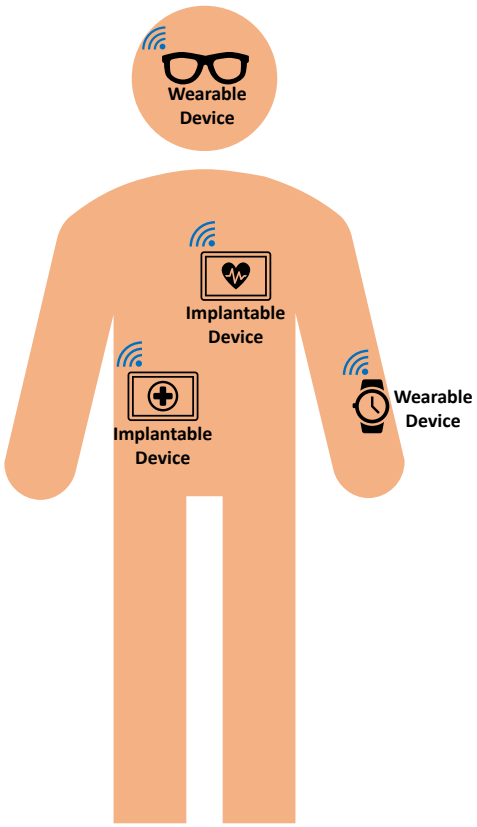

Figure 2.4: HBICS Devices Implanted and Worn on a Human Body

As introduced in Section 1.1.3 and illustrated in Figure 2.4, the focus of the HBICS concept is substantially on human wearable and implantable devices. Together with safety, security is often cited as a concern that is especially important to HBICS due to the possibility of HBICS devices being directly involved with life-critical information and potentially in hostile environments [20] [23] [24] [25]. Therefore, it is unsurprising that security is one of the most frequently discussed aspects of HBICS in the literature. As an example, adverse events to a cardiac implant device can cause heart failure [53]. This is reflected in a well-known case study in which the US Vice President Dick Cheney disabled the wireless functionality of his heart implant pacemaker due to fear of assassination through the device being hacked [26] [54] [55].

In terms of more specific security requirements, different studies cite various security attributes as applicable to HBICS. Nevertheless, *confidentiality*, *integrity* and *availability* appear to possibly be those that get discussed most often [20] [21] [26] [55] [56] [57] [58]. These security requirements can be addressed by using various security solutions. For example, confidentiality and integrity can be addressed through the use of cryptographic key generation and exchange, as well as appropriate authentication protocols [21]. As for availability, denial-of-service is known to be a major attack type in HBICS [21] [26] [53] [54] [59]. One method of carrying out such attacks is through the launch of wireless communications link jamming; to counteract this, the use of frequency hopping can be considered [24] [60].

Various security mechanisms have been proposed for HBICS. They generally cater to the unique characteristics of such systems, such as resource limitations and direct contact or close proximity to the human body. The following paragraphs briefly describe several interesting HBICS security mechanisms discussed in the literature.

Due to their inherently low resource availability, the more traditional cryptography approaches are not suitable for HBICS [24]. As an alternative, lightweight techniques, such as elliptical curve cryptography (ECC), may be used. ECC offers a higher security level, with a relatively smaller size cryptographic key. As an example, a 160-bit ECC-based key can provide an equivalent security level to a 1025-bit RSA-based key [22]. Nevertheless, ECC-based systems are also known to be more complex to build practically [57]. Another alternative approach is the use of attribute-based encryption (ABE), where messages can be encrypted for users that have a particular set of attributes [22].

Since many HBICS devices are located inside or on a human body, the use of biometrics for security purposes is often advantageous. This can be viewed as a situation in which a particular unique physiological feature of a human body is used to produce and maintain cryptographic keys [24]. Common types of features include

electrocardiogram (ECG), photoplethysmogram (PPG), fingerprint and iris [24] [26] [28] [54] [59] [61] [62]. The inter-pulse interval (IPI), or the timing between heartbeats, is perhaps one of the most prominently discussed physiological biometrics. IPI is an ECG-based mechanism and has a benefit in that it can be measured anywhere on the body. Additionally, IPI is known to have a high level of randomness, making its use as a physiological entropy source advantageous [55]. Other examples of biometrics discussed in the literature include palm images, blood pressure, blood glucose and temperature [27] [55] [61].

The fuzzy commitment scheme [63] is often discussed as being an enabler for various biometrics to be useable [28] [62] [64]. This is because biometric data are subject to random noise. Consequently, support for small variabilities in such physiological data is required. This can be catered for by the fuzzy commitment scheme, which allows for errors in what is equivalent to a decryption key to be tolerable to a certain degree [63] [65].

For intra-body communication, which is mostly applicable to Tier 1 devices, the use of nano-communication has also been discussed [26] [66] [67]. Besides the classic electromagnetic communication, nano-communication also covers the emerging field of molecular communication. This is where the human biological systems get used to encode information to biological molecules, such as proteins, which would then act as information carriers. Security-wise, biochemical cryptography is a relatively new field of study from this aspect.

There are also other interesting studies relevant to HBICS security. For example, the use of physical unclonable function (PUF) for authentication based on unique hardware attributes [61] [68], the use of game theory to counteract anti-jamming [69], the use of blockchain for authentication and distributed secure storage [21] [59], and the use of machine learning to detect attacks and to assist with authentication [28] [59].

2.3. IoT Security Solution Enabling Technologies

As previously introduced in Section 2.1, literature review has indicated that Intrusion Prevention Systems (IPS), Intrusion Detection Systems (IDS), and robust design of systems architectures and protocols are used as solutions to mitigate security threats. Enabling technologies for these solutions range from more traditional security technologies, such as encryption and digital signature, to more newly emerging technologies with a lot of active research happening, such as blockchain and machine learning.

This section is a literature review of IoT security solution enabling technologies, categorised by the focus of Chapter 3 and Chapter 4 of this thesis. Firstly, the use of blockchain is explored, in preparation for the content of Chapter 3. Subsequently, the use of machine learning is reviewed, in preparation for the content of Chapter 4.

2.3.1. Blockchain

Blockchain is one of the enabling technologies for IoT communications security solutions that is being very actively discussed in the literature. Blockchain is a secured and distributed ledger that can aid in resolving many of the problems with centralisation [70]. The term was first used by S. Haber and W.S. Stornetta in a 1991 article [30]. However, it was not until 2008 when the first blockchain system was created, which is the Bitcoin cryptocurrency system [71]. Since then, blockchain has become quite popular, with usage in various industries, such as finance, insurance, logistics, and agriculture [30].

Trust has traditionally been a barrier to the implementation of decentralised architecture. Blockchain addresses this issue through the use of distributed consensus mechanism to prove the validation of new transactions while still recognising the

earlier transaction history [72]. Popular consensus algorithms include Proof-of-Work (PoW), Proof-of-Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) [70] [73].

A few types of blockchains have been discussed in the literature, each of which has its own distinct features. The first popular one is *public blockchain* which is truly decentralised and permissionless, allowing anyone to maintain a copy of the blockchain and to validate new blocks. This is the type of blockchain that Bitcoin uses. The second one is *private blockchain* which is suitable for a single organisation usage, as it is permissioned, requiring nodes to be known members. The third type is *consortium* or *federated blockchain* which is similar to the private blockchain, but with expanded access given to multiple organisations [30].

Smart contracts can also be utilised through the use of blockchain. This is where programmable applications are stored inside the blockchain to manage and automatically execute transactions when specific terms and conditions are met. Smart contracts are supported in newer blockchain platforms such as Ethereum and Hyperledger [30].

The use of blockchain has been identified as a potential security solution for many IoT applications, including intra-vehicular and inter-vehicular networks [30] [43] [74]. Potential IoT security use cases of blockchain are such as for encryption, authentication and access control, privacy, data provenance, and integrity assurance services [70].

Nevertheless, the suitability of blockchain in IoT has also been questioned, as the system involves significant energy usage, delay and computational overhead [38] [72]. There are also other integration issues, such as security, privacy, data management, and lack of standardisation and interoperability [30]. Consequently, blockchain usage in IoT is still currently an active research area.

An alternative approach to blockchain, where directed acyclic graphs (DAG) are used, has also been discussed in the literature [30] [75]. An advantage of this approach is that scalability can be improved [30]. The consensus mechanisms used for DAG are also known to potentially be more suitable for IoT applications than traditional approaches, such as PoW and PoS, because of lower transaction fees, lower resource consumption, and the ability to achieve much higher transaction throughput [75].

2.3.2. Machine Learning

Machine Learning (ML) is another enabling technology for IoT security solutions that is being very actively discussed in the literature. Applications are such as authentication and access control, jamming attack detection, anomaly/intrusion detection, side channel leakage detection, and trust management [76] [77] [78] [79]. Apart from security, there are also many other applications that ML can be used for. Within the communications area alone, other applications include spectrum allocation, interference alignment, hardware resource allocation, link evaluation, routing path search, etc [80].

There are several ways in which ML is categorised in the literature. One way is to categorise by learning approach. When the data are labelled and the learning system is trained using such data, this is called *supervised learning*. When the learning is done through classifying unlabelled data, this is called *unsupervised learning*. The approach with algorithms that can be used on a mixture of labelled and unlabelled data is called *semi-supervised learning*. Finally, the approach in which learning is performed through rewards obtained after interacting with the environment, which is similar to the learning behaviours of humans and animals, is called *Reinforcement Learning (RL)* [76] [81] [82].

Another way of categorising ML is by whether if the learning is done through shallow learning network or large deep neural network, described using the terms *typical/basic ML* and *Deep Learning (DL)*, respectively [76] [80].

A typical/basic ML system comprises three layers. The first layer is the input layer which takes the pre-processed data as input. Subsequently, this gets passed on to the next layer, which is the feature extraction and processing layer. Within this second layer, data processing is used to extract the data patterns. Then the output layer, which is the last layer, provides the results of the given classification task [80].

With DL, however, there are multiple hidden layers between the input and output layers. This provides an advantage in allowing the data to be input to the system in raw form without requiring complex input data pre-processing. This is possible because each of the hidden layers can extract different features and progressively strengthen relevant features while weakening irrelevant features [80]. Due to the DL model being more sophisticated, it is also more effective than a typical ML model in scenarios with higher data volume [82] [83]. Nevertheless, this also comes at a cost, which is the longer training time and the higher computational power requirement [83].

There are various DL architectures found in the literature, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) [82] [84]. Another popular approach is Deep Reinforcement Learning (DRL), which is a combination of DL and RL [76] [84].

There exist several issues related to the use of ML for IoT security that are yet to be resolved, making this a very active research area. For example, the availability of suitable IoT security related datasets was identified as one of the issues [31] [76] [82]. Another issue is the potential computational and storage overheads incurred in ML/DL, which may not be suitable for IoT devices. Furthermore, the longer

convergence time of ML/DL may not be suitable for real-time applications, particularly for safety-critical systems [76].

The next issue is the susceptibility of ML/DL to misclassification and overclassification [76]. This also relates to another issue, which is the potential intentional attack to misclassify data through the use of adversarial perturbations introduced in the learning process. These perturbations may be so small that they are imperceptible to humans but can cause the system to generate harmful outputs [6] [76] [85] [86].

Several studies such as those by Chaabouni et al. [31], Al-Garadi et al. [82] and Hussain et al. [76] have also identified issues related to the applicability of current knowledge and how to keep knowledge up to date. One issue is that the model trained to solve one problem may not perform well in solving a slightly different problem in a similar field. For example, a trained model may have difficulty detecting unknown and zero-day attacks. In addition, there is a need to develop schemes for managing knowledge transfer and the continuous learning of new threats.

2.4. The Future of IoT Security: Quantum Technology

Quantum technology is known to be revolutionising communications systems and networks in the near future [87]. Quantum communications and computing can provide strong security through the use of quantum keys based on the quantum no-cloning theorem and uncertainty principle. The security comes from the physics of quantum science, in which if eavesdroppers carry out observations, measurements or copy actions, the quantum state will be disturbed, and consequently, their actions would be easily detected. Quantum computing will also provide an enabling platform for artificial intelligence applications requiring big data and massive training [88]. This would potentially aid in resolving some ML timing and performance issues identified in Section 2.3.2.

On the downside, quantum technology will also cause many functions of current communications security systems to be significantly degraded. For example, Sliwa [89] stated that quantum computing can break public cryptography and will consequently disable all electronic payment systems, potentially destroying the economy. The Elliptic Curve Digital Signature Algorithm (ECDSA), which is widely used in blockchain systems, is also known to be vulnerable to quantum computing attacks due to the elliptic curve digital logarithm problem not being a hard problem for quantum computing [71]. Consequently, quantum computer processing power may enable adversaries with sufficient resources to execute majority attacks to take over certain blockchain systems in the future [74].

Major research efforts in the field of quantum technology are currently in progress [90]. Quantum computing and communications have been identified as security and long-distance networking enabling technologies for research in 6G communications systems and beyond [88] [90]. The Quantum Internet, which is a network envisioned to connect quantum devices with classical ones [91], is also being researched.

Chapter 3 Blockchain for Vehicular Privacy Enhancement

Research areas relating to connected and autonomous vehicles (CAVs) are currently of high interest to both academia and industry. This is not surprising as transportation is one of the essential aspects that people around the world interact with on a daily basis. With the increasing communications signal transmissions from vehicles, comes an increasing risk of those vehicles being easily tracked. This issue has triggered the formation of a significant research area in vehicular location privacy, which is still very active. The advancement of blockchain development, over the past decade or so, triggers it to become a potential Internet of Things (IoT) security solution enabling technology that is worthy of investigation. In this chapter, the use of a consortium blockchain system to enhance vehicular privacy in a multi-jurisdictional road network is explored.

3.1. Introduction

Connected and autonomous vehicles (CAVs) are known to form the backbone of future intelligent transportation systems (ITS) to provide travel comfort, road safety, and other value-adding services [6]. It is therefore unsurprising to find that vehicular ad hoc networks (VANETs) have become a very active topic of discussion in both industry and academia.

Security is described in the literature as being of high importance in VANETs because connected vehicles are directly related to road safety. A compromised system could therefore directly result in potential accidents, injuries, and possibly casualties [92].

Moreover, security, including privacy, is known to be a challenge in VANETs because of contributing factors such as vehicles being highly mobile, making it difficult to ensure security and non-repudiation [40] [41] [47]. Furthermore, identification information included in the messages sent out by vehicles, such as a public key certificate, together with the associated radiofrequency emission, can be used to derive the identity of a particular vehicle's operator and track their exact location information at a particular time [92]. Consequently, vehicular location privacy is a topic that has been actively discussed in the literature. A feasible approach for mitigating location privacy threats is the use of pseudonyms instead of real vehicle identifications. Nevertheless, there is also a trade-off for this with non-repudiation. A solution to this is to make the use of vehicular pseudonyms conditional, allowing real identity to be traceable by authorised parties [92].

Roadside units (RSUs) are stationary infrastructure usually deployed on the roadside or at other designated locations, such as intersections or carparks [93]. They are known to be typical components of VANETs [94]. There is an expectation that owing to the ease of deployment and inexpensive price of wireless technology, RSUs will be abundantly available on roadsides to provide wireless access to vehicles [95]. This seems even more likely given the potential increase in spectrum availability for short-range communications from technologies such as millimeter-wave (mmWave) and visible light communication (VLC), which are emerging trends in vehicular communications [10].

As will be discussed further in Section 3.2, the use of Public Key Infrastructure (PKI) for pseudonymous communications is known to be highly viable. Indeed, PKI is also known to be flexible and is well recognised as being reliable for use in trust establishment owing to its evolution over a time span of more than twenty years [96]. For the above reasons, it is possible that economically practical pseudonym schemes used in the near future will take advantage of the mature PKI technology and the wide

availability of RSUs. Such schemes would need to be designed in a security-focused manner.

Blockchain is known to have an underlying technological architecture similar to that of PKI; however, authentication and nonrepudiation cannot be guaranteed reliably. Conversely, blockchain can be used together with PKI to eradicate weaknesses and form a more strengthened system [96].

As will be further discussed in Section 3.2, blockchain is an area that is still being very actively researched and developed, owing to the fact that it is a relatively new technology. Since its inception, the definition of blockchain has shifted away from being associated with cryptocurrency to much broader associations covering numerous applications. Such evolution also comes with changes to the blockchain's fundamentals. For instance, there have been developments in alternatives to the use of the original but inefficient Proof-of-Work (PoW) consensus algorithm. Newer types of blockchains, such as the permissioned and the consortium types, help in this regard, as they limit access to only authorised and authenticated users, thereby removing the requirement for PoW.

The evolution of blockchain to become more efficient, flexible, and scalable, together with its distinct immutable distributed ledger property, has been identified as a potential solution to assist in the management of conditionally anonymised vehicular pseudonyms. This is especially so for the consortium blockchain type, which has great potential in the provision of a secured integrated solution for a road network jointly managed by different jurisdictions (e.g., a national road network made up of portions managed by different states). A literature review conducted has identified no existing research on this particular topic. However, there appear to be some limited existing works that discuss the use of more traditional blockchain architectures to improve vehicular pseudonym allocation and issuance management, which also have other shortcomings that should be addressed.

For the above reasons, this chapter proposes a novel scheme of PKI vehicular pseudonym issuance and management system by taking advantage of the forecasted abundant presence of RSUs and the consortium permissioned blockchain system, a high-level conceptual operating model of which is depicted in Figure 3.1. As will be further explained in Section 3.3, the proposed scheme also addresses shortfalls identified in the very limited existing research works briefly discussed above.

The remainder of this chapter is organised as follows. Section 3.2 outlines existing related works, including a preliminary overview of vehicular pseudonym change approaches and of blockchain as an enabling technology for security, decentralisation, and collaboration. Section 3.3 captures the motivations and contributions of the proposed scheme. The architecture of the proposed framework is then presented in Section 3.4. Section 3.5 discusses the simulation environment used for the proposed scheme. This follows by Section 3.6 which discusses the simulation results and performance analysis. Finally, conclusion is given in Section 3.7.

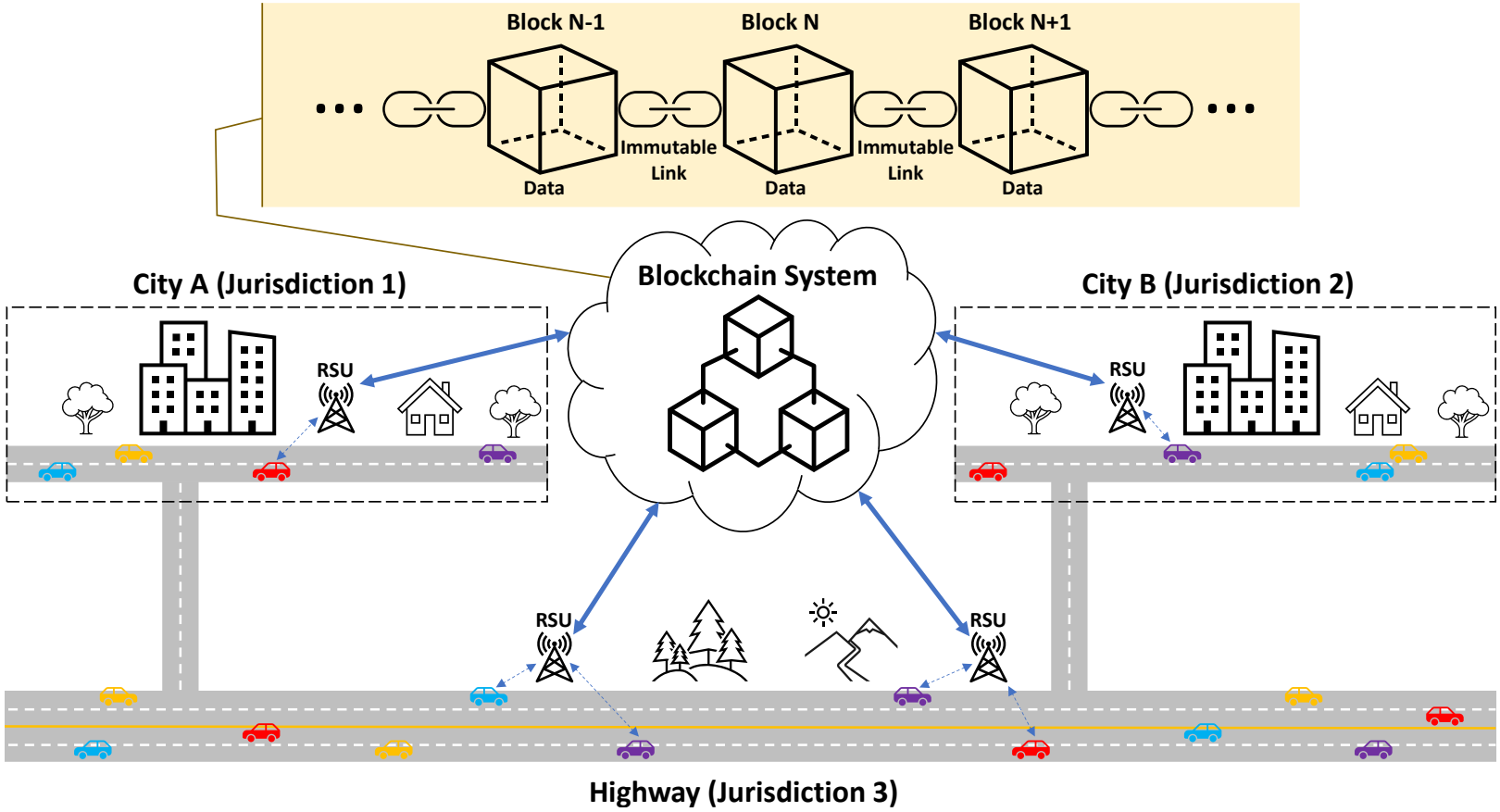

Figure 3.1: Conceptual Operating Model of the Proposed Scheme

3.2. Related Works

3.2.1. Vehicular Pseudonym Change Approaches

The use of vehicular pseudonyms works on the principle that if vehicles attach long-term identities to outgoing messages, they could potentially be tracked and linked using a trace of received messages and the associated location information [34]. Therefore, for pseudonym schemes to be effective in mitigating linkability, the pseudonym needs to be frequently changed. The European standard ETSI TS 102 867 recommends that pseudonyms are changed every five minutes, whereas the American standard SAE J2735 recommends that this is done every 120 s or 1 km, whichever occurs last [97] [98].

3.2.1.1. Pseudonym Schemes

A number of *privacy-preserving authentication* schemes are discussed in the literature, also often referred to interchangeably as *pseudonym* schemes. These schemes can be classified into five categories: *Public Key Infrastructure (PKI)* or *asymmetric cryptography* schemes, *identity-based cryptography* schemes, *certificateless cryptography* schemes, *symmetric cryptography* schemes, and *group signature* schemes. Interested readers can refer to existing survey papers such as [92] and [93] for a more detailed analysis of these.

This chapter focuses on the PKI category, which is known to be one of the most viable approaches, and is what early vehicular networks privacy-preserving authentication proposals were based on, after which many major initiatives also followed [92]. In this category, vehicles use public key certificates issued by a trusted certification authority

(CA) as a means of authentication. Communication is achieved by attaching a public key certificate and the corresponding digital signature to any message being sent out. The certificate and signature can then be verified for authenticity by message receivers [93].

To preserve communication privacy, vehicles would use their public key certificates and the associated public/private key pair in a short-term manner, essentially operating a public key certificate as a pseudonym.

3.2.1.2. Pseudonym Lifecycle

Vehicular pseudonyms undergo the following three life cycle phases: *pseudonym issuance*, *pseudonym use,* and *pseudonym change*. In addition, there are additional optional phases of *pseudonym resolution* where the real identities of misbehaving vehicles are determined from their pseudonyms, and *pseudonym revocation* where the pseudonyms of misbehaving vehicles are revoked. As can be seen, in order for relevant authorities, such as law enforcement agencies, to be able to identify misbehaving vehicles through their pseudonyms, it is essential that the privacy protection mechanisms of such vehicular pseudonym schemes only allow for conditional anonymity and not complete anonymity [92].

There are two major approaches of pseudonym issuance discussed in the literature, namely, third-party issuance and self-issuance. The third-party issuance approach involves an external pseudonym issuing authority to create pseudonyms for vehicles. In some schemes, vehicles would request short-term pseudonyms in certain intervals, while in others, pseudonyms would be pre-loaded in large amounts sufficient to last up to a few years. The self-issuance approach was introduced to reduce communication overhead with the CA by enabling vehicles to generate pseudonyms themselves [92] [93].

The possibility of vehicles having multiple pseudonyms available for use at a given time introduces the risk of potential Sybil attacks [34]. This is when an adversary claims to be multiple vehicles with different identities and acts maliciously to mislead other vehicles to make harmful decisions [7]. Self-issuance approaches are known to be more difficult to avoid Sybil attacks because of the higher level of autonomy of vehicles [92]. In contrast, a controlled on-demand approach to pseudonym issuance that can limit the number of valid pseudonyms available to each vehicle at a time would potentially offer higher protection against Sybil attacks [34]; however, this also means that vehicles must request new pseudonyms at certain intervals, which may introduce scalability issues [92]. The revocation of pseudonym certificates is known to be another scalability challenge, as the verification of pseudonyms against a large certificate revocation list (CRL) may be impractical because of the associated high computational cost [93].

3.2.1.3. Support from Roadside Units (RSU)

As RSUs are known to be a typical component of VANETs [94], it is unsurprising to find them being used to perform different tasks in pseudonym schemes. The roles of RSU in existing pseudonym schemes include the issuance of pseudonyms, the management of group keys, and pseudonyms verification [92].

3.2.2. Blockchain: Security, Decentralisation and Collaboration

Blockchain is known to have the potential to constitute security solutions, including in the areas of access control, data integrity, confidentiality, and availability. This partially emerged from the blockchain's inherent security properties in cryptocurrency networks that have been shown to be capable of mitigating attacks such as distributed denial-of-service (DDoS) attacks, modification attacks, and double

spending [30]. Furthermore, the use of blockchain has been recommended to replace centralised cloud platforms, which are susceptible to single point of failure issues around security and privacy [99]. The use of blockchain has also been identified more specifically as a potential security solution for many Internet of Things (IoT) applications, including VANETs [30] [43] [100].

At a higher level, blockchain is an enabling technology for various solution domains that is being actively discussed in the literature. It is a secured and distributed immutable ledger that can assist in rectifying many problems related to centralisation [70]. The term was originally used by Haber and Stornetta in their article from 1991 [30]. However, it was only as recently as 2008, when the first blockchain system was created, being the Bitcoin cryptocurrency system [71]. Blockchain has since become quite popular, with its usage in numerous other industries, such as insurance, logistics, and agriculture [30].

The issue of trust has traditionally been an obstacle to the implementation of a decentralised architecture. Blockchain addresses this issue by using distributed consensus mechanism that recognises earlier transaction history when validating new transactions [72]. Well-known consensus algorithms include Proof-of-Work (PoW), Proof-of-Stake (PoS) and Practical Byzantine Fault Tolerance (PBFT) [70] [73].

Smart contracts can also be used with blockchain. This refers to programmable applications stored inside the blockchain to manage and automatically carry out transactions in situations where specific terms and conditions are met. Turing-complete smart contracts are supported in newer blockchain platforms, including Hyperledger and Ethereum [101] [102].

Several types of blockchains exist in the literature, each of which has its own unique features. Firstly, one popular type is the *public blockchain*, which Bitcoin uses. This type is truly permissionless and decentralised, allowing anyone to maintain a copy of the blockchain and validate new blocks. Next is the *private blockchain,* which is

permissioned, requiring nodes to be known members. This blockchain type is thus suitable for single-organisation use. Similar to the private blockchain, but with expanded access to multiple organisations is the *consortium* or *federated blockchain*. This blockchain type allows auditable, reliable, and synchronised distributed database recording of data exchanges between participating members [30]. Consortium blockchain, therefore, has the potential to improve the effectiveness of collaboration between different organisations.

3.2.2.1. Hyperledger Fabric Blockchain Platform

Hyperledger is an open-source production of cross-industry blockchain technologies introduced by the Linux Foundation in 2016 [100] [103]. It consists of multiple blockchain platform projects, a result of the adherence to the "no one fits all" concept. These platforms include Fabric, Iroha, and Sawtooth [104].

Hyperledger Fabric was the first proposal of Hyperledger projects [104]. The platform is known to be multi-purpose [99], has permissioned access control [104], and can operate across multiple different organisations with private data support that can also be exchanged between members of a subset of a consortium [103].

In terms of the support for private data, the official documentation [105] states that Hyperledger Fabric breaks this down into two different elements, namely the actual private data and a hash of the private data. The actual private data are sent to authorised organisations in a peer-to-peer manner using the gossip protocol. It is only the hash of such private data that gets submitted to the consensus services and stored in the blockchain's blocks. This mechanism allows the immutability verification of private data while disallowing unauthorised organisations from accessing it even when they have access to the blocks.

As mentioned earlier, Hyperledger Fabric also has smart contract feature, which the platform calls “chaincodes”. Furthermore, the platform is known to have low latency, with the ability to process up to 10000 tps and record a blockchain transaction in approximately 0.5 s, even with peer nodes located in different continents [104]. For these reasons, this platform was selected for use in a simulation of the proposed scheme, as detailed in Sections 3.5 and 3.6.

The Hyperledger framework proposes the use of the Crash Fault Tolerant (CFT) ordering service as its consensus protocol, which offers faster speed and higher scalability than the more common PoW and PoS algorithms [73] [100]. Although this may provide lower security against malicious or faulty nodes, it is not so much an issue because the framework is a permissioned blockchain, and consequently, the ability to modify its ledger can be controlled by only allowing trusted parties [100].

The CFT protocol works by having a leader ordering service node getting dynamically elected, and the followers adhering to the leader’s decisions. It is known to be capable of tolerating failures up to 50% [103].

3.2.3. Blockchain for Vehicular Pseudonym Management

As outlined in Section 3.1, existing works that discuss the use of more traditional blockchain architecture to improve vehicular pseudonym allocation and issuance management are quite limited. Bao et al. [98] and Benarous et al. [106] appear to be the only relevant existing works in this area. A summary of these and their shortfalls are provided below.

In [98], a pseudonym shuffling scheme is proposed, based on the authors previous work on blockchain-based dynamic key management [107]. Their proposed scheme works by having existing PKI-based pseudonyms shuffled and reused by different vehicles where the shuffling allocation between different vehicles is done through the

RSUs acting as access points. The RSUs form groups, each of which is managed by a privacy manager based on geographical distribution. Shuffling results are recorded in a blockchain that uses the original but inefficient PoW consensus mechanism. The scheme allows for the flexibility to minimise the storage space and potential CRL size if each vehicle only carries a few allocated pseudonyms at a time; however, this would require very frequent contacts with RSUs, which may be impractical in locations such as remote areas.

In [106], a blockchain-based pseudonym management framework is proposed, where pseudonym generation is performed purely by vehicles without interference by authorities. The main part of the framework consists of two blockchains for storing certified pseudonyms and revoked pseudonyms, managed by registered vehicles and RSUs, respectively. However, in order to authenticate any received message, a vehicle must also check the two blockchains to ensure that the pseudonym used has actually been certified and has not been revoked. Therefore, for this scheme to function effectively, the participating vehicles would need near real-time access to the blockchains, which may be impractical in reality. Furthermore, scalability of the blockchain system may also be an issue, as it would need to support read and write access from all participating vehicles, and would need to allow vehicles to access its ledger's most recent state in a very timely manner. The paper also did not include any simulation results to confirm the feasibility of the proposed scheme.

It is important to note that there are other works in the literature that discuss the use of blockchain with authentication, admission, and revocation in VANETs, such as [108], [109], [110] and [111]. However, the focus of these works is not on vehicular pseudonym allocation and issuance management, although some partial overlaps may exist because of the common properties of blockchain.

3.3. Motivations and Contributions

As introduced in previous sections, the use of pseudonyms instead of real vehicle identifications is a known mitigation for location privacy preservation. Nevertheless, there is also a trade-off for location privacy, which is non-repudiation. To resolve such a trade-off, the use of vehicular pseudonyms needs to be made conditional, so that the real identity is allowed to be traced by authorised parties. The use of blockchain systems has a potential to assist in the management of conditionally anonymised vehicular pseudonyms due to the associated distinct immutable distributed ledger property. However, as can be seen from the discussions in Section 3.2.3, none of the existing related works have focused on vehicular pseudonym issuance and management in a multi-jurisdictional road network, and in particular, the complexities in data handling at interfaces between these different jurisdictions and the associated security risks. These shortfalls and the existence of consortium blockchain systems as a potential solution enabling technology, thus, motivate the development of a novel pseudonym issuance and management scheme proposed in this chapter. The proposed scheme uses permissioned consortium blockchain paired with the traditional PKI-based cryptography system to carry out pseudonym issuance and management in a dynamic, secure, conditional privacy-preserving and distributed manner, while also enabling integrated collaboration between different organisations.

Table 3.1 summarises the comparison of the proposed scheme to the existing works derived from the discussions in Section 3.2.3 and from the analysis results in Section 3.6.

Table 3.1: Comparison of the Proposed Scheme to Existing Works

	Bao et al. [98]	**Benarous et al.** [106]	**Proposed Scheme**
Conditional Privacy Support	Yes	Yes	Yes

	Bao et al. [98]	**Benarous et al.** [106]	**Proposed Scheme**
Conditional Privacy Support for Multiple Jurisdictions Road Network	Not discussed. Likely to require additional high complexity for secured access and management.	Not discussed. Likely to require additional high complexity for secured access and management.	Fundamental integrated secured access and management support through consortium blockchain architecture.
Blockchain Consensus Mechanism	Proof of Work	Proof of Elapsed Time for the certifying block. Round Robin for the revocation block.	Scheme's architecture is flexible. The efficient CFT protocol used in the simulation.
Connectivity to Blockchain Requirements	Through very frequent contacts with RSUs – every 40-50 minutes based on scenario presented in their performance analysis.	Vehicles need near real-time access to the blockchains for scheme to function effectively.	Through intermittent contacts with RSUs – once a week or more frequent.
Vehicle's Storage Requirements	Low to moderate – only 10 pseudonyms required to be stored at a time based on scenario presented in their performance analysis but likely to require more for the scheme to be practical.	Potentially very high – entire blockchain records may need to be stored if near real time remote access to the blockchains network cannot be provided.	Moderate – consisting of longer term default pseudonyms (indicative guide of 100 pseudonyms) plus dynamically assigned pseudonyms depending on vehicle's usage time (in the order of 10s to 100s for average daily vehicle usage of 11 hours or less).
Scheme's Critical Time Obstruction Point	Pseudonym revocation verification by infrastructure – CRL size is low based on scenario presented in their performance analysis but likely to be much higher for the scheme to be practical.	Vehicle requires near real-time access to the blockchains to perform pseudonym verification and revocation checks in order for the scheme to function effectively.	Pseudonym revocation verification by infrastructure – CRL size potentially much smaller than [98]. No requirements for vehicle to have near real-time access to the blockchains.

	Bao et al. [98]	**Benarous et al.** [106]	**Proposed Scheme**
Scheme's Feasibility Demonstration	Simulation carried out.	No records of any simulation carried out.	Simulation carried out.

3.4. System Architecture

In this section, the system architecture is proposed, as shown in Figure 3.2, in which the components and their integration are further elaborated in Section 3.4.1. The architecture considers a scenario in which vehicles are allowed to move freely throughout the entire road network, but different areas within the network are managed by N different organisations. An example of this is a national road network in a country, which is made up of roads managed by different jurisdictions (e.g., different states).

Due to the number of different stakeholders involved, interface management between different jurisdictions can become complex. A permissioned consortium blockchain technology is proposed to be used to mitigate such complexity and enable better collaboration between jurisdictions, while simultaneously ensuring high security and conditional anonymity in vehicular pseudonym use.

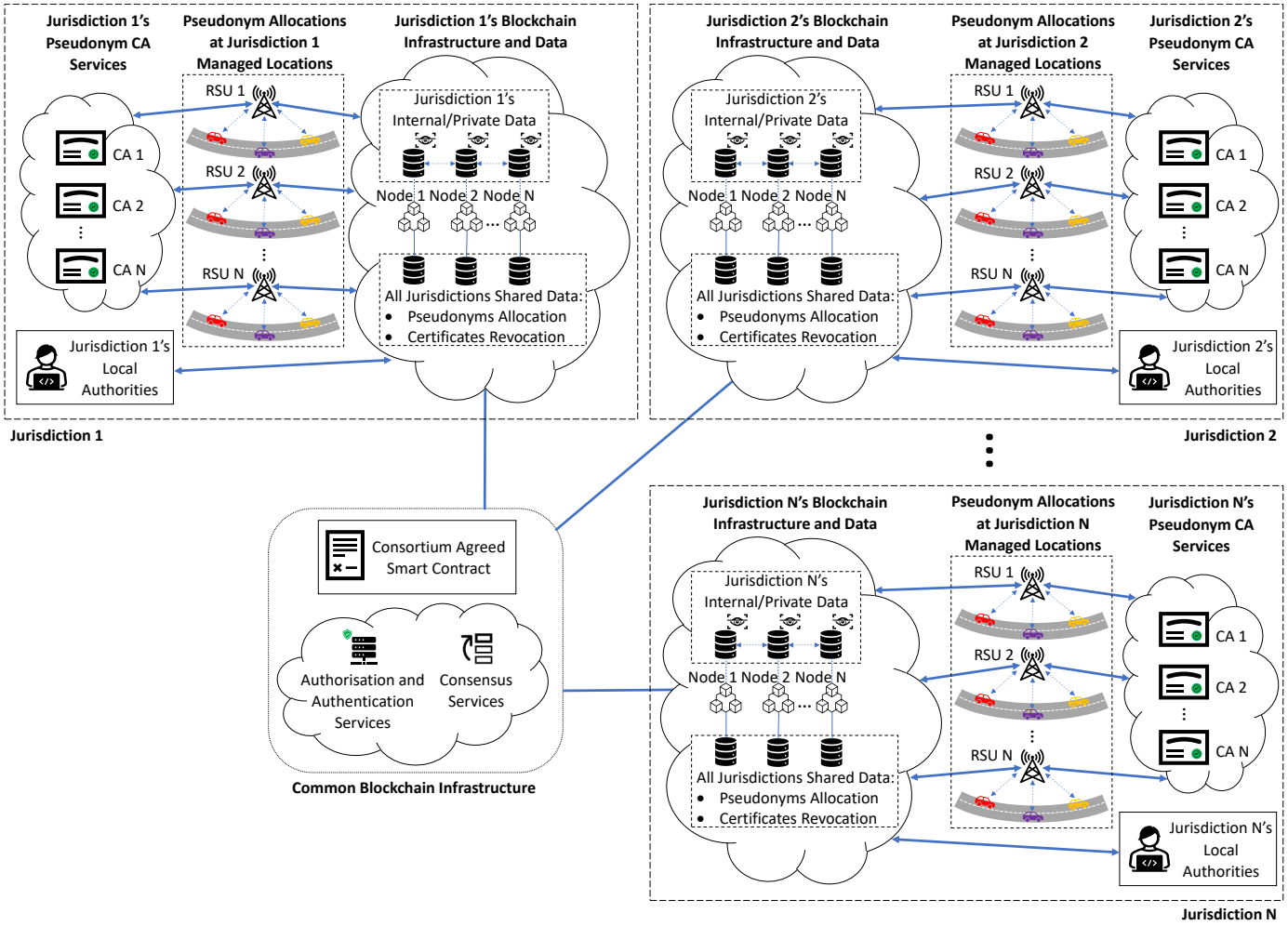

Figure 3.2: System Architecture

3.4.1. System Components

The proposed system consists of several main components, the first of which is a consortium *blockchain network* supporting multiple jurisdictions. The blockchain network used in this architecture is to have permissioned access control, which allows different jurisdictions to collaborate, yet still maintain their independence. Core elements in the network include an agreed smart contract, consensus services, authorisation and authentication services, and blockchain nodes managed by each local jurisdiction. The blockchain network also includes an immutable ledger for storing vehicular pseudonym issuance records and the CRL, which is consistently distributed to all jurisdictions and accessible by their local authorised users (e.g., local law enforcement agencies). At the same time, any jurisdiction's internal data associated with the issuance of pseudonyms can also be recorded but made inaccessible to other jurisdictions.

Other major components of the system include *RSUs*, *pseudonym CA services*, and *vehicles*. The RSUs act as an interface between the blockchain network, pseudonym CA services, and vehicles. They liaise with the vehicle and the CA in carrying out pseudonym issuance tasks and record relevant information into the blockchain. As for CAs, their function is for issuing PKI pseudonym certificates. They are to be managed by each local jurisdiction. They can be cloud-based and can be shared among other jurisdictions if required. Vehicles, which are the end users of issued pseudonyms, interact with the system by requesting pseudonym issuance through RSUs located at different roadside locations throughout different jurisdictions.

Lastly, *local authorities*, such as the government department responsible for the management of transportation or law enforcement agencies also play a major role in the system architecture. They access the blockchain through their jurisdiction's

blockchain nodes. Access can be used for purposes such as incident investigations and pseudonym certificate revocations.

3.4.2. System Functionality Description

The following sub-sections detail the functionality of the proposed system at each phase of the pseudonym lifecycle defined in [92].

3.4.2.1. Pseudonym Issuance

In the proposed scheme, a typical vehicle V_i is preloaded with a set of default pseudonyms P_{default}. This set of pseudonyms is expected to be used infrequently when no valid pseudonyms allocated by the proposed scheme are available. The pseudonyms in P_{default} are to have relatively long expiry dates, such as in the order of years, so that they do not have to be replaced too often. To reduce the ability of V_i from potentially carrying out Sybil attacks on others and to discourage routine usage of pseudonyms in this set, the allocation of P_{default} needs to be limited.

V_i's pseudonyms set for regular usage P_{usage} would get issued in a more dynamic manner as illustrated in Figure 3.3. Pseudonyms in this set are to have short expiries, in the order of days or even hours. As pointed out earlier, this helps mitigate vehicles misusing issued pseudonyms to carry out Sybil attacks. Furthermore, this also helps to minimise the potential volume of revoked certificates required to be in the CRL.

Pseudonym issuance occurs when V_i travels into a serviceable area of RSU_j which is a participant RSU in the scheme. In such a situation, V_i would use its pseudonym $p_x \in P_{\text{default}} \cup P_{\text{usage}}$ to identify itself and send a Certificate Signing Request CSR_y, which contains the public key for its proposed new pseudonym certificate p_y, to RSU_j. Thereafter, RSU_j would then verify PKI certificate p_x for validity and check against the

CRL using the blockchain's smart contract to ensure p_x has not been revoked. If all goes well, it then passes on the request to its associated pseudonym issuance authority CA_j. Once the issuance of the new pseudonym p_y by CA_j has been carried out, RSU_j would transmit it to V_i and uses the blockchain's smart contract to record the transaction details in a format pre-agreed with the other jurisdictions. The transaction record would at least capture details of p_x and p_y, and may also include the jurisdiction's internal data that are to be logged but made inaccessible to other jurisdictions.

Note that similar to any typical PKI system, the trust establishment between V_i and RSU_j is done through the verification of each other's public key certificate (i.e., the pseudonym p_x in the case of V_i) to ensure non-expiry and issuance by a trusted CA. In addition, note that the transmissions of CSR_y and p_y between V_i and RSU_j need to be encrypted to prevent any potential eavesdroppers from being able to link p_y to p_x. With additional overhead, these communications can also be digitally signed by the sender to help protect against other security threats such as spoofing. An example of this is where hypothetically a malicious vehicle V_m pretends to be the owner of pseudonym p_x which it actually took from an earlier message sent to it by V_i. In such a situation, V_m might try to use p_x to launch an attack by falsely identifying itself and sending CSRs to waste RSU_j's and CA_j's resources on useless tasks. Such a threat would be mitigated if V_m is required to digitally sign the request with the private key associated with p_x it does not have. Nevertheless, this is not functionally necessary otherwise because when RSU_j transmits p_y encrypted using the public key associated with p_x, the vehicle V_m would still not be able to decrypt due to not having the private key.

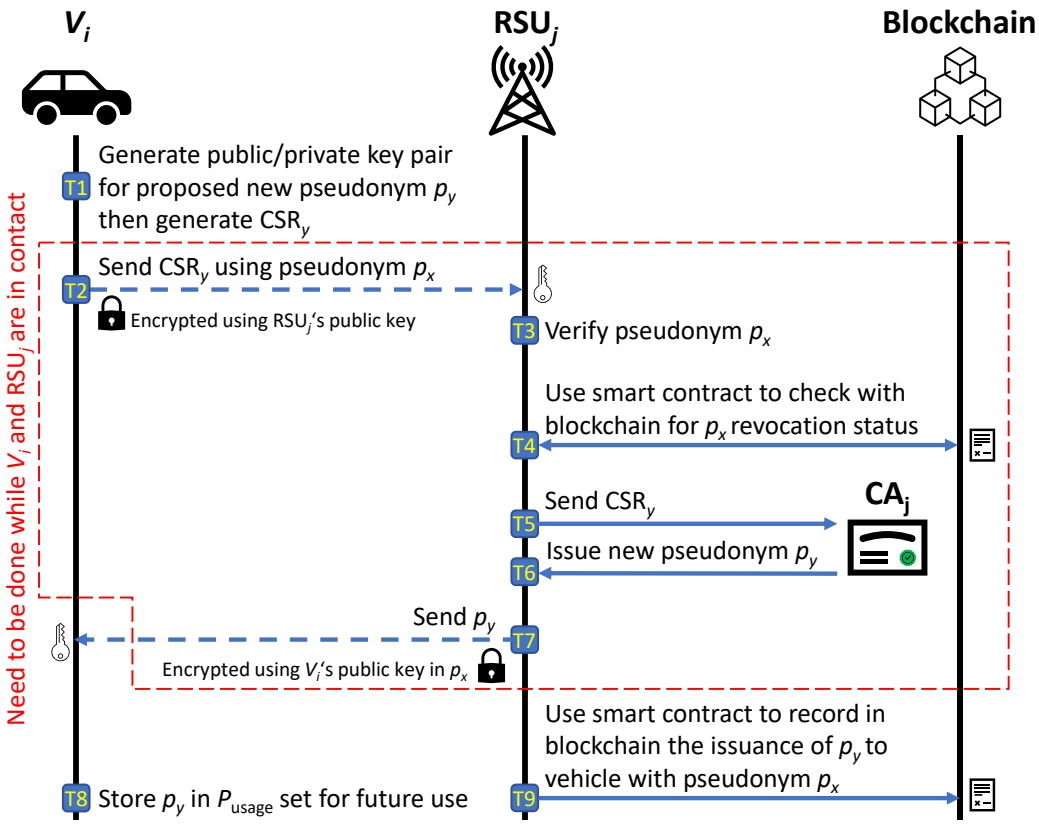

Figure 3.3: Pseudonym Issuance Process

3.4.2.2. Pseudonym Use

In this phase, a vehicle V_i would select a pseudonym $p_y \in P_{\text{usage}}$, which is one of the pseudonyms issued to it from the previous phase, to sign its outgoing messages. The recipient vehicle V_r would then be able to assess the credibility of the message by carrying out two common tasks associated with the use of PKI in any order. First, it needs to verify the signature against the actual message content and the public key in p_y. Second, it needs to verify that p_y has not expired and was issued by a trusted CA. Note that in this scheme, it is assumed that PKI certificates for all the CAs used are already preloaded to the vehicles, and so are readily available for verification tasks.

For practical reasons, the proposed scheme assumes that V_r does not have access to the blockchain or CRL to verify that p_y has not been revoked. However, the need for this has already been mitigated by making p_y having short expiry and by having V_i's previous pseudonym used when requesting for p_y checked against the CRL before p_y's issuance. Therefore, the likelihood of p_y having already been revoked at any given time is relatively low.

Of note, however, is the fact that if V_i has exhausted all the pseudonyms available in P_{usage} set, it may have to fall back on using $p_y \in P_{\text{default}}$, which has long expiry. In the case where V_r receives a message with such p_y attached, there would be a higher likelihood that p_y may have already been revoked, and therefore, V_r should give V_i a reduced trustworthiness rating and treatment.

The trustworthiness ratings can thus be categorised as shown in Table 3.2.

Table 3.2: Pseudonym Use Trustworthiness Rating Categorisation

Trustworthiness Category	**Criteria**
Highly Trusted	$p_y \in P_{\text{usage}}$ used (valid pseudonym with short expiry period of hours or days)
Partially Trusted	$p_y \in P_{\text{default}}$ used (valid pseudonym with long expiry period of a few years)
Not Trusted	Invalid/expired pseudonym

3.4.2.3. Pseudonym Change

Technical details on the determination of the exact situations when a vehicle would change its pseudonym are outside the scope of this study. There have been many pseudonym change schemes discussed in the literature during the past few years that can possibly be matched with the pseudonym issuance scheme proposed in this study. Interested readers may refer to existing survey papers such as [49], which provides a comprehensive survey of pseudonym change strategies based on a mix-zone where

vehicles change their pseudonyms on the basis of predefined locations, and a mix-context where vehicles independently determine the time and location to change their pseudonyms.

3.4.2.4. Pseudonym Resolution

As the proposed scheme ensures conditional anonymity rather than complete anonymity, pseudonym resolution by authorised personnel can be carried out. Resolutions can be performed using the blockchain's smart contract to trace the pseudonym issuance history of a particular vehicle backward or forward.

High data integrity in this process can also be expected due to the immutability property of blockchain resulting from the distributed consensus mechanism that recognises the earlier transaction history. Figure 3.4 illustrates how this is achieved based on the implementation found in the Hyperledger Fabric platform, which records the cryptographic hash of the previous block in the header of each block.

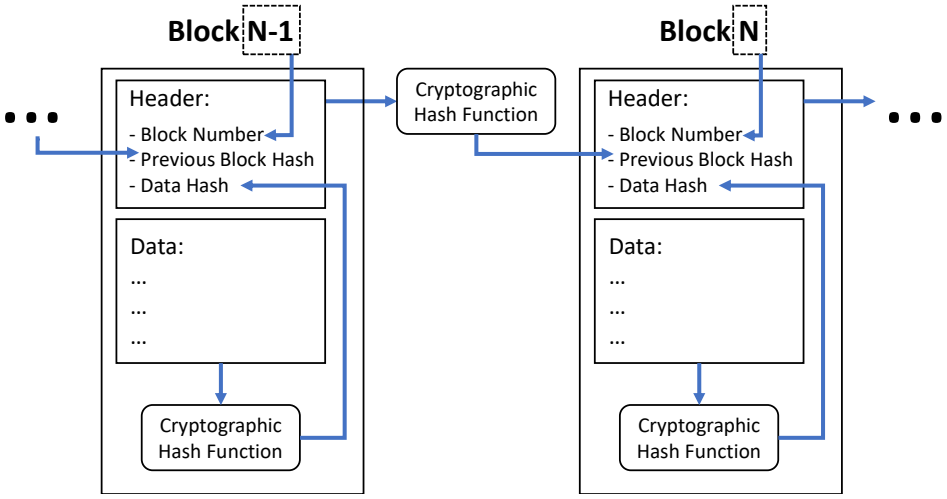

Figure 3.4: Hyperledger Fabric Blocks Linkages

3.4.2.5. Pseudonym Revocation

Pseudonym revocation is performed by adding revoked pseudonyms to the CRL captured in the blockchain ledger. This task is performed by authorised personnel of any jurisdiction by invoking a smart contract function. If required, the smart contract function can also be programmed to trace the pseudonym issuance history of a particular vehicle backward and forward to revoke all unexpired pseudonyms.

As pseudonyms issued by the proposed system have short expiry, they do not need to stay in the CRL for a long time. The smart contract can be used to assist in automating the deletion of expired pseudonyms from the CRL. The ability to minimise the size of the CRL is quite important as it can alleviate the time taken in the pseudonym revocation checking process undertaken during pseudonym issuance as discussed previously.

The process of pseudonym revocation and removal of expired entries is illustrated in Figure 3.5.

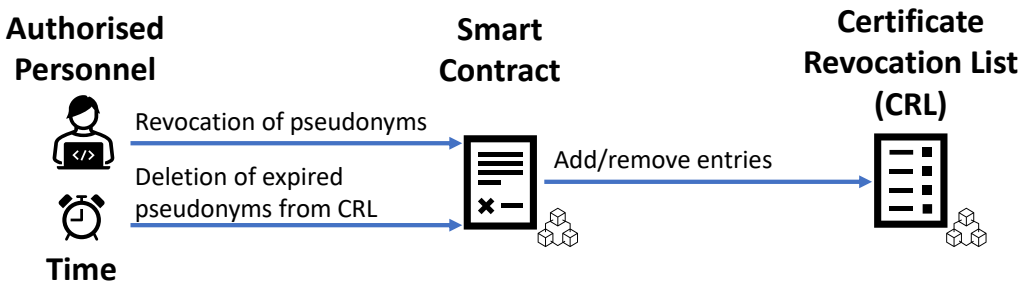

Figure 3.5: Pseudonyms Revocation and Removal of Expired Entries

3.4.3. Blockchain Design

Here, the design of key components of the blockchain system used in the proposed architecture is discussed.

3.4.3.1. Ledger

The blockchain ledger consists of two mandatory types of records with the attributes as shown in Figure 3.6. The first is the IssuanceRecord type, which captures pseudonym issuance transactions. The RSUID attribute is used for storing transaction identification information by the managing RSU; this may include information such as the RSU's identification, and the date and time of the transaction. The vehicle's current pseudonym p_x is stored in the CurrentPseudonym attribute, and the new pseudonym p_y is stored in the NewPseudonymIssued attribute.

The second type is the RevocationRecord which captures each pseudonym that has been revoked. The records of this type collectively form the CRL of the scheme. The AuthorityID attribute is used for storing transaction identification information; this may include information such as the authorised party's identification, and the date and time of the transaction. The pseudonym being revoked is stored in the PseudonymRevoked attribute. The PseudonymExpiry attribute captures the extracted expiry date and time information of the pseudonym being revoked. This attribute facilitates the automated deletion of expired pseudonyms from the CRL by the smart contract.

It is important to note that RSUID and AuthorityID do not necessarily capture the immutable transaction information like what is stored in the actual blockchain's blocks. Thus, the information stored in these attributes should not be solely relied upon for security critical auditing tasks.

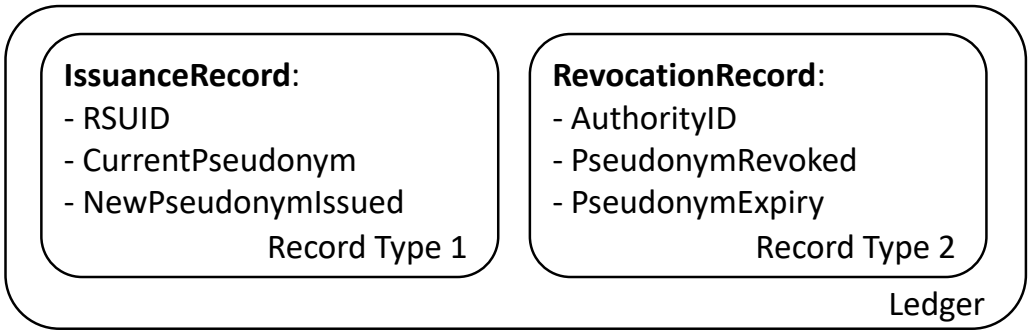

Figure 3.6: Blockchain's Ledger – Record Types

3.4.3.2. Smart Contract

As shown in Figure 3.3 and Figure 3.5, several processes require the invocation of smart contracts. These processes are pseudonym revocation checks, pseudonym issuances, pseudonym revocations, and expired pseudonym deletions. Algorithm 3.1 through to Algorithm 3.4 provide details on how these processes are to be implemented in the smart contract. In addition, Algorithm 3.5 can also be implemented to supplement Algorithm 3.2 if support for private data accessible only internally within each jurisdiction is required. Note that the implementation of Algorithm 3.5 is based on how private data works in Hyperledger Fabric and may need to be adapted if a different blockchain platform is used.

Algorithm 3.1: Check Pseudonym Revocation Status

```
Input: Pseudonym p
1  begin
2      for each ledger record r of type
       RevocationRecord do
3          if r.PseudonymRevoked = p do
4              return true
5          end if
6      end for
7      return false
8  end
```

Algorithm 3.2: Record Pseudonym Issuance

Input: RSU's transaction identification information i_{RSU},
Vehicle's current pseudonym p_x,
Newly issued pseudonym p_y

begin
  Create record r of type IssuanceRecord
    Set $r.\text{RSUID} = i_{\text{RSU}}$
    Set $r.\text{CurrentPseudonym} = p_x$
    Set $r.\text{NewPseudonymIssued} = p_y$
  Submit r for processing to update ledger and add to blockchain
end

Algorithm 3.3: Revoke Pseudonym

Input: Authority's transaction identification information i_a,
Pseudonym p,
Pseudonym expiry date and time e

begin
  Create record r of type RevocationRecord
    Set $r.\text{AuthorityID} = i_a$
    Set $r.\text{PseudonymRevoked} = p$
    Set $r.\text{PseudonymExpiry} = e$
  Submit r for processing to update ledger and add to blockchain
end

Algorithm 3.4: Remove Expired Revoked Pseudonym from Ledger

Input: Current date and time d

begin
  for each ledger record r of type RevocationRecord **do**
    if $r.\text{PseudonymExpiry} < d$ **do**
      Delete r from ledger
    end if
  end for
end

Algorithm 3.5: Record Jurisdiction's Private Data Associated with Pseudonym Issuance

Input:	Private data identification information i_v, RSU's transaction identification information i_{RSU}, Private data v
1	**begin**
2	Submit i_v, i_{RSU} and v for processing to update private ledger and add hash of private data to blockchain
3	**end**

3.4.3.3. Consensus Mechanism

The proposed scheme does not prescribe the use of any particular consensus algorithm as long as such an algorithm is supported by the consortium blockchain platform being used. Nevertheless, in order to optimise the performance, and because the consortium blockchain is permissioned, it is envisaged that inefficient consensus algorithms required for traditional permissionless blockchain (e.g., PoW) would not be used.

Note that the CFT consensus protocol is used by the Hyperledger Fabric consortium blockchain platform in the simulation discussed in Section 3.5.

3.4.3.4. Pseudonym Issuance Cost

Based on tasks T1 to T9 illustrated in Figure 3.3, the communication and computation time costs associated with the proposed architecture can be derived as detailed in Table 3.3.

Table 3.3: Pseudonym Issuance Communication and Computation Time Costs

Task	Communication Time Cost	Computation Time Cost
T1	N/A	t_{T1P1}: Generate private key for p_y
		t_{T1P2}: Generate public key for p_y
		t_{T1P3}: Generate CSR_y
T2	t_{T2M1}: Over the air message capturing CSR_y and current pseudonym p_x	t_{T2P1}: V_i encrypts message
		t_{T2P2}: RSU_j decrypts message
T3 (fixed cost)	N/A	t_{T3P1}: PKI verification of V_i's current pseudonym p_x
T4 (fixed cost)	t_{T4M1}: Message to the blockchain's network capturing p_x	t_{T4P1}: Execute smart contract to check for p_x's revocation status
	t_{T4M2}: Message from the blockchain's network capturing p_x's revocation status	
T5	t_{T5M1}: Message capturing CSR_y	N/A
T6	t_{T6M1}: Message capturing p_y	t_{T6P1}: Issue new pseudonym p_y
T7	t_{T7M1}: Over the air message capturing p_y	t_{T7P1}: RSU_j encrypts message
		t_{T7P2}: V_i decrypts message
T8	N/A	N/A
T9	t_{T9M1}: Message to the blockchain's network capturing p_y and p_x	t_{T9P1}: Execute smart contract to record that new pseudonym p_y issued to vehicle with pseudonym p_x

The total time cost for pseudonym issuance to a vehicle can be summarised as $t_{\mathrm{total}} = t_{\mathrm{fixed}} + t_{\mathrm{variable}}$, where t_{fixed} is the fixed cost regardless of the number of pseudonyms a vehicle requests issuance for, and t_{variable} is the variable cost that varies depending on the quantity q of pseudonyms being requested. From Table 3.3, t_{fixed} and t_{variable} can be defined as follows:

$$t_{\mathrm{fixed}} = t_{\mathrm{T3P1}} + t_{\mathrm{T4M1}} + t_{\mathrm{T4M2}} + t_{\mathrm{T4P1}} \quad (1)$$

$$t_{\mathrm{variable}} = q \times (t_{\mathrm{T1P1}} + t_{\mathrm{T1P2}} + t_{\mathrm{T1P3}} + t_{\mathrm{T2M1}} + t_{\mathrm{T2P1}} + t_{\mathrm{T2P2}} + t_{\mathrm{T5M1}} + t_{\mathrm{T6M1}} + t_{\mathrm{T6P1}} + t_{\mathrm{T7M1}} + t_{\mathrm{T7P1}} + t_{\mathrm{T7P2}} + t_{\mathrm{T9M1}} + t_{\mathrm{T9P1}}) \quad (2)$$

3.5. Simulation Environment

A simulation of the proposed scheme was set up on a desktop computer with an Intel i7 2.90 GHz processor, 32 GB of random access memory (RAM), and Windows 10 Enterprise operating system. The integrated traffic and network simulator Vehicles in Network Simulation (Veins) [112] Version 5.0 was used to pair up with the blockchain platform Hyperledger Fabric [113] Version 2.3. Note that Veins integrates the traffic simulator SUMO [114] (Version 1.2.0 used in the experiment) with the network simulator OMNeT++ [115] (Version 5.5.1 used in the experiment). In addition, OpenSSL [116] Version 1.1.1i was also integrated into the simulation for the provision of PKI services. The interactions between Veins and Hyperledger Fabric, and between Veins and OpenSSL were performed through the command line interface built into OMNeT++. Note that version selections of software used were based on the latest stable and compatible releases available at the time.

The Veins framework was chosen partly because it uses SUMO mobility simulator, which has high-performance simulation capability for large networks, and is known to be more suitable for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications than alternative mobility simulators such as METACOR and VanetMobiSim. Another reason is that Veins uses the OMNeT++ network simulator, which is known to be a flexible tool for researchers to use in high-mobility VANETs applications. Veins enables SUMO and OMNeT++ to work with each other in an integrated manner [93].

As already introduced in Section 3.2.2.1, Hyperledger Fabric is a feature-rich permissioned consortium blockchain platform, which indicates that it is suitable for the implementation of the proposed system architecture. In addition to what was already discussed, Hyperledger Fabric's official documentation [117] states that versions 2.x of the platform offer several enhancements to the "private data" feature.

Consequently, this could also make the Hyperledger Fabric platform more suitable for use with the proposed scheme in terms of confidentiality in working with jurisdiction's internal data, as detailed in the architecture outlined in Section 3.4.

A small-scale version of the proposed architecture was successfully simulated, as shown in Figure 3.7. The simulation setup took advantage of the predefined test network included with the Hyperledger Fabric platform Version 2.3. The test network has two predefined organisations, each of which has one peer node.

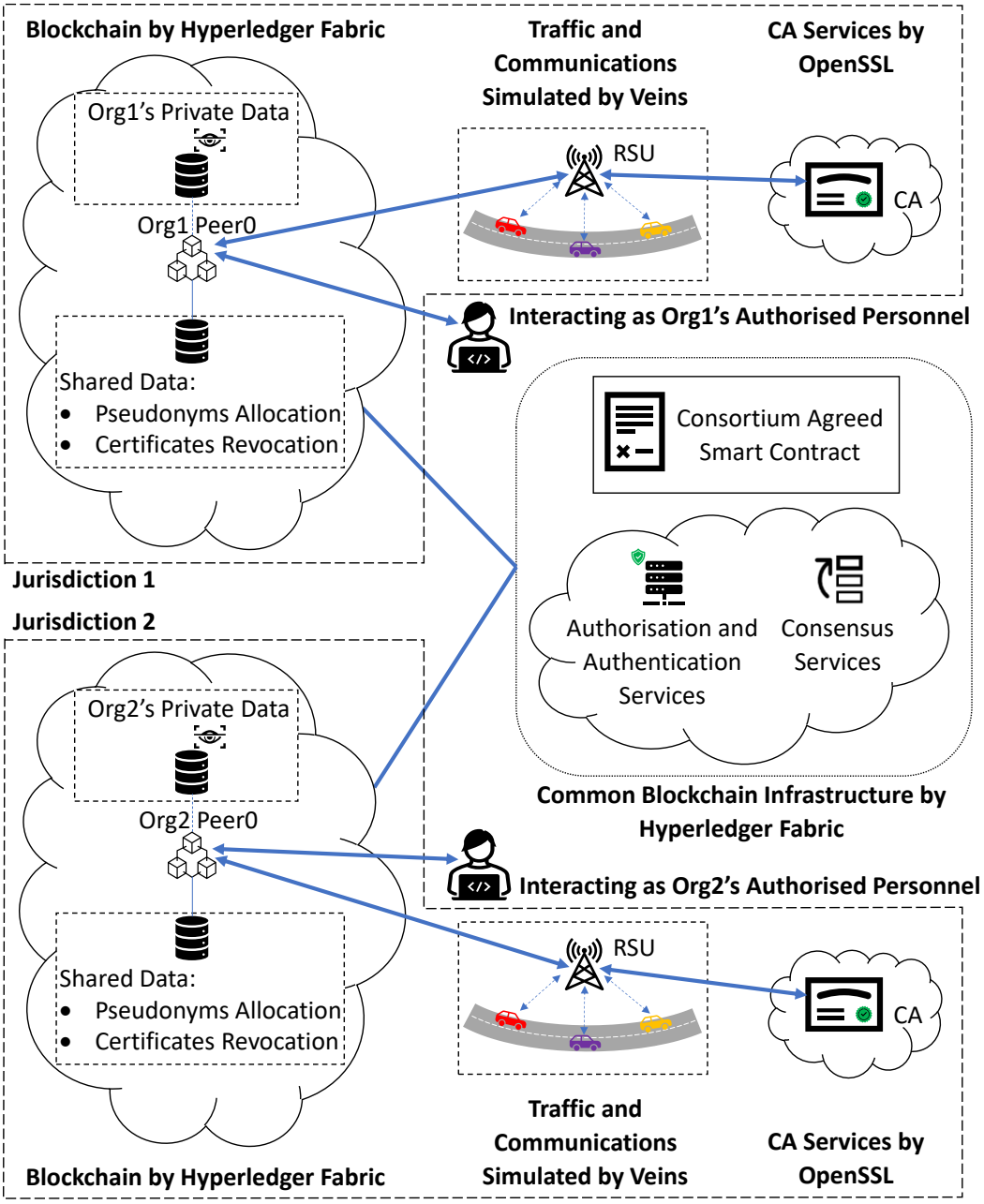

Figure 3.7: Simulated System Architecture

The pseudonym allocation process shown in Figure 3.3 was implemented to test the concurrent pseudonym allocations of two RSUs, which represent two different locations managed by two different jurisdictions. Ten vehicles were generated to drive into each of the service areas at the arrival rate of one vehicle every five seconds. In addition, the use of smart contract to keep the CRL short, as shown in Figure 3.5, was also tested by using time to trigger an automation process in deleting expired pseudonyms from the list.

From the blockchain's point of view, Algorithm 3.1 through to Algorithm 3.4 were deployed onto the smart contract. In addition, the optional Algorithm 3.5 was also deployed to experiment with the private data feature of the blockchain platform. The smart contract was written in JavaScript, which is one of the three languages supported by the platform, the other two being Go and Java. The smart contract functions were called from JavaScript applications using an application programming interface (API). Subsequent to the simulation, verification tasks were carried out through the use of smart contract to look up the jurisdictions shared data recorded in the blockchain, as well as each jurisdiction's private data permitted to be accessed only by its own personnel.

From the key management and encryption side, 2048-bit RSA asymmetric encryption was used for both vehicular pseudonyms and RSU's public key certificate. Owing to file size limitation, encryption of messages communicated between vehicles and the RSU was performed using 256-bit AES symmetric encryption, with such symmetric keys exchanged securely between the vehicle and the RSU through asymmetric encryption.

3.6. Simulation Results and Evaluation

3.6.1. Key Simulation Results

In general, all functionalities of the integrated simulation system were found to be working as expected, and the results demonstrate the feasibility of practical deployment of the proposed scheme. Nevertheless, analysis can be done more closely on the performance results, focusing on tasks bordered within the red dashed box shown in Figure 3.3, which are the time critical tasks that can only be carried out while the vehicle is in contact with the RSU. Table 3.4 shows the average execution time results of these tasks after ten runs. Note that the results exclude the infrastructure-to-infrastructure communication costs t_{T4M1}, t_{T4M2}, t_{T5M1} and t_{T6M1} because of the assumption that they are negligibly low, and thus were not controlled as part of the simulation setup. It is also important to note that it would be possible for the RSU to distribute tasks such as t_{T2P2}, t_{T6P1} and t_{T7P1} to be executed in parallel by different processes, which would improve the performance in terms of the required contact time between the vehicle and RSU. Subjecting to collective demand at an RSU, this parallel processing may even be deemed essential to cater for cases where a large number of pseudonyms are requested by a vehicle for issuance.

Table 3.4: Average Execution Time of Critical Tasks

Cost Notation (Refer Table 3.3)	**Task Description**	**Average Execution Time (ms)**	
t_{T2P1}	V_i: Encrypt message containing CSR_y	62.9	
t_{T2M1}	V_i: Over the air transmission of message containing CSR_y and current pseudonym p_x to RSU_j	0.2	Note: artificially simulated time as obtained from Veins
t_{T2P2}	RSU_j: Decrypt message containing CSR_y	114.7	
t_{T3P1}	RSU_j: Current vehicle's pseudonym p_x PKI verification	53.9	

Cost Notation (Refer Table 3.3)	Task Description	Average Execution Time (ms)	
t_{T4P1}	RSU_j: Check pseudonym p_x revocation status with blockchain	15.4	when CRL size = 10
		16.7	when CRL size = 50
		17.7	when CRL size = 100
		36.5	when CRL size = 500
		54	when CRL size = 1000
		188.8	when CRL size = 5000
		356.9	when CRL size = 10000
		1719.1	when CRL size = 50000
		3421.9	when CRL size = 100000
t_{T6P1}	CA_j: Issue new pseudonym p_y	63.9	
t_{T7P1}	RSU_j: Encrypt message containing p_y	121.9	
t_{T7M1}	RSU_j: Over the air transmission of message containing p_y to V_i	0.3	Note: artificially simulated time as obtained from Veins

From the results, it can be seen that the task of checking pseudonym p_x revocation status depends greatly on the blockchain's CRL size at any given time. For this reason, a few different CRL sizes were experimented with. To provide a better illustration, a graph is plotted, as shown in Figure 3.8. Indeed, the time taken to carry out a pseudonym revocation status check against CRL has the potential to impose a significant delay in the order of multiple seconds if the CRL size is sufficiently large. When compared to the other tasks, this delay can potentially be seen as the most significant performance bottleneck of the scheme. This is not surprising, as it aligns with the literature, such as [93], as previously discussed. Therefore, it is crucial to keep the CRL size as small as possible.

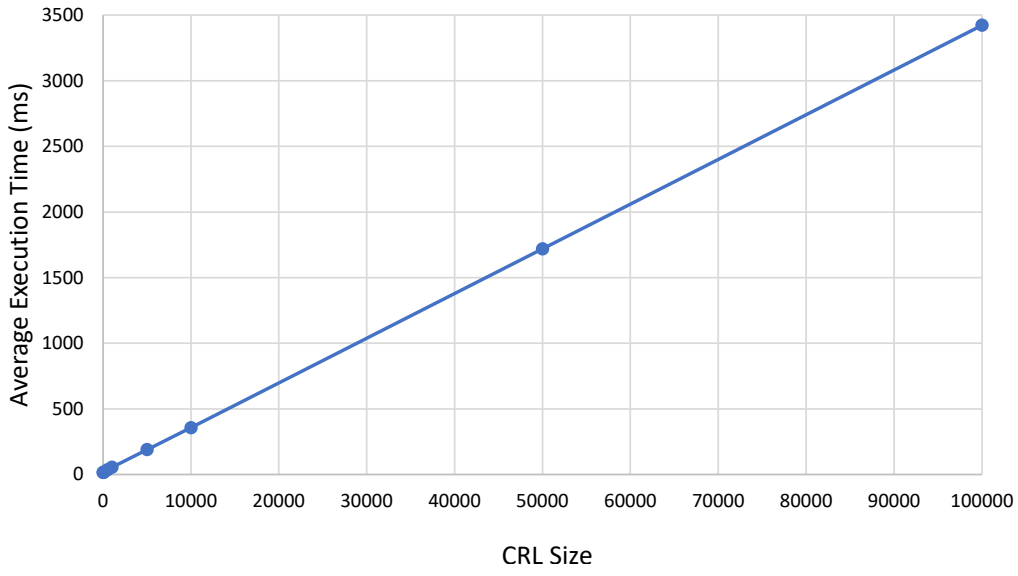

Figure 3.8: Average Pseudonym Revocation Status Check Time – Different CRL Sizes

3.6.2. Performance Comparison with Existing Methods

As outlined in Section 3.3, none of the existing schemes in the literature seems to fully cover the functionality scope of the proposed pseudonym issuance and management scheme discussed in this chapter. Nevertheless, readers may find that performance comparisons with the existing schemes at the subsystem level, as presented in this sub-section, are still of value. Note that Table 3.1 can also be referred to for summary.

3.6.2.1. Data Handling at Organisational Interface Comparison

It is well known in systems engineering and management that interfaces can introduce complexities and associated problems and risks [118] [119]. Similarly, in the scenario discussed in this chapter, interfaces between different jurisdictions would introduce complexities in data handling. For example, the same piece of policy could be interpreted differently by each jurisdiction and thus get inconsistently applied. This brings in security risks, such as in the areas of access control, data integrity,

confidentiality, and availability. Existing common methods of mitigating these risks, especially on an interorganisational basis, do not form an integrated solution. Consequently, various other risks, including further security risks introduced through human errors, may be exacerbated. Existing solutions, such as [98] and [106] outlined in Section 3.3, left the interorganisational interfacing aspect predominantly unaddressed; therefore, some or all of these security risks would still remain. In contrast, the proposed scheme mitigates these risks using a permissioned consortium blockchain system. For instance, access control is mitigated by the permissioned blockchain's authorisation and authentication services. Data integrity is mitigated by the distributed consensus mechanism, which recognises the earlier transaction history, and the use of smart contracts that ensure intended functionalities are consistently implemented. Jurisdiction's internal confidentiality is also honoured by allowing the integrated use of jurisdictions private data. Finally, the use of a consortium blockchain promotes availability by connecting different jurisdictions together in a distributed manner, with multiple blockchain nodes providing redundancy.

3.6.2.2. Storage Requirement Comparison

As mentioned in Section 3.2.1.2, some existing pseudonym schemes favour pseudonyms being pre-loaded in large amounts sufficient to last up to a few years [92]. Assuming the European standard ETSI TS 102 867 is followed where pseudonyms are changed every five minutes [97] [98], a vehicle would require 12 pseudonyms per one hour usage. For a vehicle that is used for an average of x hours a day, in a year (365 days), such a car would require $12 \times 365x = 4380x$ pseudonyms if none of them are to be reused. If the car is to be pre-loaded with pseudonyms to last for y years, the amount would grow linearly, requiring $4380xy$ pseudonyms in total. In contrast, the amount in the scheme proposed in this chapter would remain constant

regardless of y, requiring only a fraction of pseudonyms to be stored at any given time. For example, assuming pseudonyms have a maximum validity duration of one week, the maximum storage required at any one time would only be for $7x \times 12 = 84x$ representing a potential storage space saving of over $52y$ times less space. Note that for simplicity, the small number of default pseudonyms pre-loaded to the proposed scheme's vehicles are excluded in this analysis.

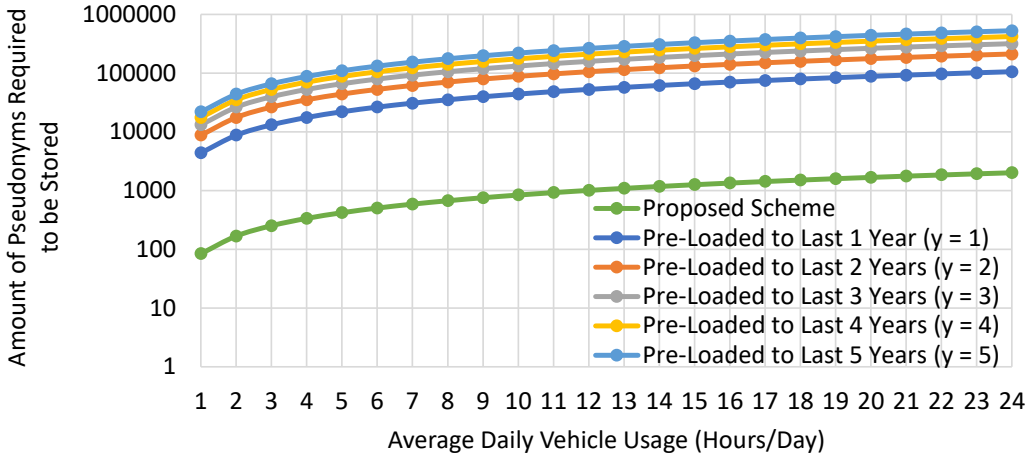

Figure 3.9: Amount of Pseudonyms Required to Be Stored per Vehicle Comparison – Proposed Scheme vs. Pre-Loading Pseudonyms

Figure 3.9 illustrates the comparison of the number of pseudonyms required to be stored per vehicle, based on a given average daily vehicle usage hours, for the scheme proposed in this study versus what would otherwise be required for pre-loading pseudonyms to last from one to five years. Table 3.5 represents this in terms of the percentage of additional storage space that would otherwise be required for pre-loading pseudonyms relative to the scheme proposed in this study.

Table 3.5: Storage Space Percentage Comparison Assuming Four Hours per Day Average Daily Vehicle Usage – Proposed Scheme vs. Pre-Loading Pseudonyms

Number of Years Pseudonyms Pre-Loaded to Last For	Pseudonyms Required – Proposed	Pseudonyms Required – Pre-Loaded	Percentage of More Storage Space Required Relative to the Proposed Scheme
1	336	17520	5214.29%
2	336	35040	10428.57%
3	336	52560	15642.86%
4	336	70080	20857.14%
5	336	87600	26071.43%

3.6.2.3. CRL Size Comparison

As outlined in Section 3.2.1.2 and demonstrated in the simulation carried out, a large CRL is undesirable as it increases the computational cost. Because the proposed scheme uses pseudonyms with short expiry and a blockchain system that supports smart contracts, it is possible to minimise the size of the CRL by removing pseudonym entries that have already expired. Borrowing the example from Section 3.6.2.2, in order to revoke all pseudonyms of one car in the proposed scheme, the CRL would only be required to hold $84x$ entries plus a limited number of default pseudonyms z which is to be a very small amount; an indicative value might be $z = 100$. More importantly, after no more than a week, most of these pseudonyms can be removed from the CRL, leaving only z entries there for a longer term. In contrast, revoking a car pre-loaded with pseudonyms to last for y years would require $4380xy$ entries. Assuming that one car is revoked per day, in a year, the CRL would have $4380xy \times 365 = 1598700xy$ entries. This is significantly higher than the CRL in the proposed scheme, which would only be at approximately $(84x \times 7) + (z \times 365) = 588x + 365z$ entries. Figure 3.10 illustrates this graphically based on a given average daily vehicle usage hours, where the proposed scheme has a default pseudonym amount of 100 versus pre-loading pseudonyms to last for one and two years.

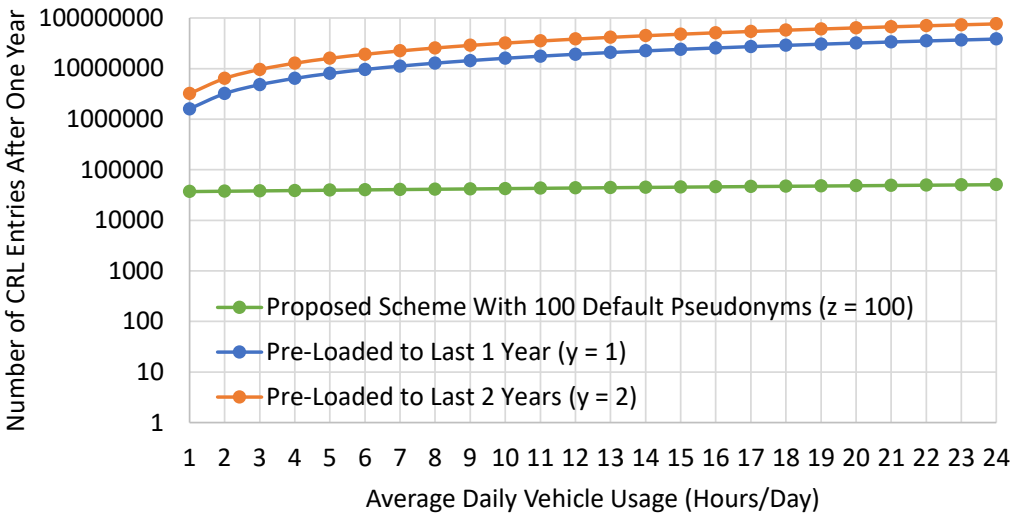

Figure 3.10: CRL Entries Comparison – One Vehicle Revoked per Day After a Year – Proposed Scheme vs. Pre-Loading Pseudonyms

Other more recent schemes discussed in the literature, such as [98], anticipated the issue with a large CRL and designed their schemes to minimise the CRL size. According to the scenario described in the performance analysis discussion of [98], where each vehicle carries only 10 pseudonyms at a time, in order to revoke such a vehicle, only 10 entries would need to be added to the CRL long-term. Although this would result in slightly smaller CRL size than the scheme proposed in this study, there is a significant trade-off. For the scenario to work, vehicles would need to establish communication with the RSU to shuffle pseudonyms every 40-50 minutes if they are to change pseudonyms every five minutes, as per the European standard ETSI TS 102 867. However, such frequent contacts with RSUs may be impractical in reality at many locations, such as in remote areas. A more practical scenario might be for vehicles to hold enough pseudonyms to last for a few days, meaning that each vehicle may need to hold hundreds of pseudonyms instead. If such a situation occurs, the CRL size of their proposed scheme would potentially become much larger than that of the scheme proposed in this chapter. In contrast, the example discussed here for the scheme proposed in this chapter, where the pseudonym validity period can be up to a week, has already taken this potential infrequent contact with RSUs into account.

Furthermore, the CRL size of the proposed scheme can be further improved by reducing the number of default pseudonyms z from 100 to a lower amount; however, this would result in a trade-off in terms of location privacy protection in situations where the default pseudonyms need to be used. Figure 3.11 illustrates the comparison of CRL sizes after a year of one vehicle being revoked per day where an average daily vehicle usage is chosen to be four hours per day ($x = 4$) as an example.

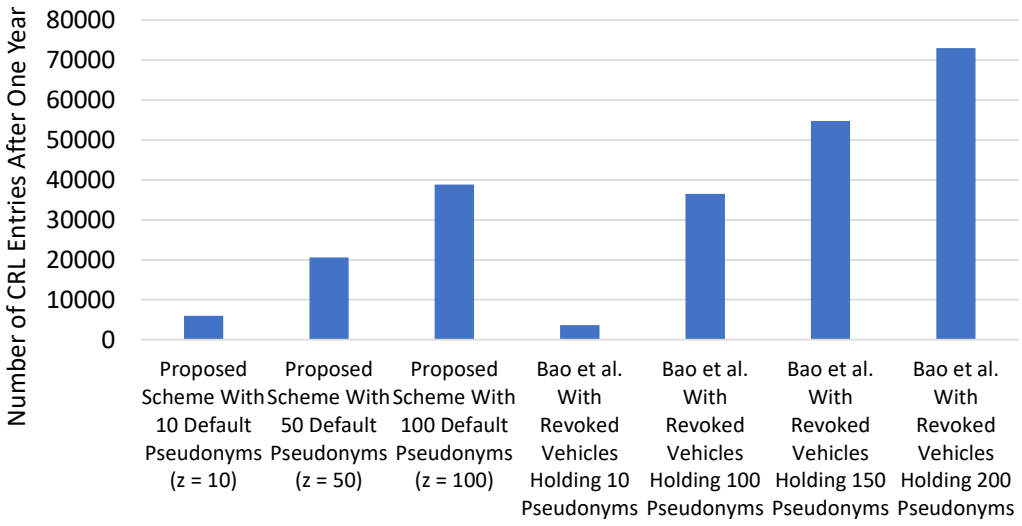

Figure 3.11: CRL Entries Comparison – One Vehicle Revoked per Day After a Year – Proposed Scheme (Four Hours per Day Average Daily Vehicle Usage) vs. Bao et al. [98]

3.6.2.4. Blockchain Consensus Algorithm Efficiency Comparison

As introduced in Section 3.2.2, there are various distributed consensus mechanisms to validate blockchain transactions. The solution outlined in [98] uses Proof-of-Work (PoW) which is the original mechanism used in Bitcoin [120]. The authors of [98] cite PoW's successful test usage relative to other consensus mechanisms as the reason for their choice. However, because PoW requires nodes to compete in solving a cryptographically hard problem, the associated unnecessarily high computational resource consumption makes it undesirable for many applications [75]. This is

especially true for situations where the use of permissioned blockchain systems is suitable, as it can eliminate the need for such expensive proofs, making consensus services a lot more straightforward [30]. The architecture proposed in this chapter uses a permissioned blockchain system, allowing for a higher efficiency consensus. Furthermore, the rapid development of blockchain solutions in recent years has created more suitable option candidates that have been successfully tested in reality through various applications. For instance, the Hyperledger Fabric platform used in the simulation here, which employs the more resource-efficient CFT consensus protocol called RAFT, is known to be one such successful solution in the industrial and IoT domains [103].

3.6.3. Additional Considerations Prior to Deployment

Although simulation results demonstrate the feasibility of practical deployment of the proposed scheme, there are still issues that may require further investigation. This is especially in relation to the variability in demand for pseudonym issuance, especially their peaks. For example, the demand could be time-dependent and varies throughout a given day, week, month, or year at a specific RSU location. It is only when the collective demands throughout the network can be characterised that an accurate expected system load can be determined. This could greatly affect the design of the system, such as the determination of the number of peer nodes in each organisation and consensus service redundancy.

Also related is the blockchain transactions processing throughput limitation; for example, as discussed in Section 3.2.2.1, in the case of Hyperledger Fabric, this is known to be up to 10000 tps. Although none of the tasks within the proposed scheme requiring such processing (e.g., task T9 shown in Figure 3.3) were identified as time critical, it might still be important to study more closely how the system would behave

if such a limit is reached, to ensure the system's robustness. Indeed, task T9 can probably afford some delay in processing, but cannot afford to be completely ignored.

Another related point to note is the observation made in Section 3.6.1 regarding the potential requirement for parallel execution of RSU's tasks by different processes. The detailed determination of such an arrangement will also need to be studied further once the expected time-dependent demand for the RSU at a particular location of interest can be quantified.

Finally, it is important to note that there may be other potential risks associated with the selected permissioned consortium blockchain platform. This is especially true given that blockchain is a technology that is still under very active research and development. Therefore, it is important to thoroughly identify and appropriately mitigate these potential risks prior to deployment.

3.7. Conclusion

In this chapter, a scheme is proposed to facilitate secure and conditional privacy-preserving vehicular pseudonym issuance and management in a multi-jurisdictional road network. The proposed architecture takes advantage of the increasingly mature permissioned consortium blockchain technology, the predicted wide availability of RSUs, and the highly viable, flexible, and well-established PKI technology. A small-scale simulation of the proposed architecture was successfully carried out using the Veins platform for integrated traffic and network simulation services, the Hyperledger Fabric platform for blockchain services, and the OpenSSL platform for PKI services. Simulation results demonstrate the feasibility of practical deployment of the scheme and identify further performance optimisation issues that should be investigated once the time-dependent demand data at different RSU locations throughout the network are obtained. In terms of comparison with existing works, performance analysis has

revealed that the proposed scheme addresses the identified shortfalls, including the ability to achieve a better balance between connectivity and storage requirements.

Chapter 4 Machine Learning for Sybil Attack Detection in the Internet of Flying Things

Although unmanned aerial vehicles (UAVs) have been used in the military domain for many decades, it was only more recently that UAVs are gaining increasing usage in civilian applications. This trend, together with the UAVs integration with the Internet of Things (IoT) to form the Internet of Flying Things (IoFT), make the IoFT becoming a very active and significant area that is gaining a lot of research attention. Just like many other IoT research areas, the topic of IoFT security is still considered challenging, with a lot of research efforts being exercised. Having been a well-known security threat to the IoT, the Sybil attack is also known as a threat to the IoFT, but relatively little research works have been carried out on it so far. On the other hand, machine learning (ML) is increasingly being used in the literature to address various challenges, including IoT security. Thus, this artificial intelligence tool has a strong potential to be suitable in being used for Sybil attack detection in IoFT, as explored in this chapter.

4.1. Introduction

Unmanned aerial vehicles (UAVs), also known as drones, refer to pilotless aerial vehicles that are either autonomously controlled by a computer or remotely controlled by a pilot on the ground. UAVs deployment in the military domain dates back several decades, with the primary applications being strike, reconnaissance and border surveillance. However, more recently, UAVs have also gained increasing usage in civilian applications, including search and rescue operations, environmental sensing and monitoring, and delivery of food and other products. In this context, the flying ad hoc network (FANET) paradigm, which is a subclass of mobile ad hoc network (MANET)

where the nodes possess aviation characteristics, is strongly tied to the operation of UAVs due to the needs for UAV nodes to communicate with each other or with other node types, such as ground control station and satellite. Consequently, the FANET paradigm and its integration with the Internet of Things (IoT) to form the Internet of Flying Things (IoFT), as depicted in Figure 4.1, have been gaining increased attention in the research community [15] [16] [121].

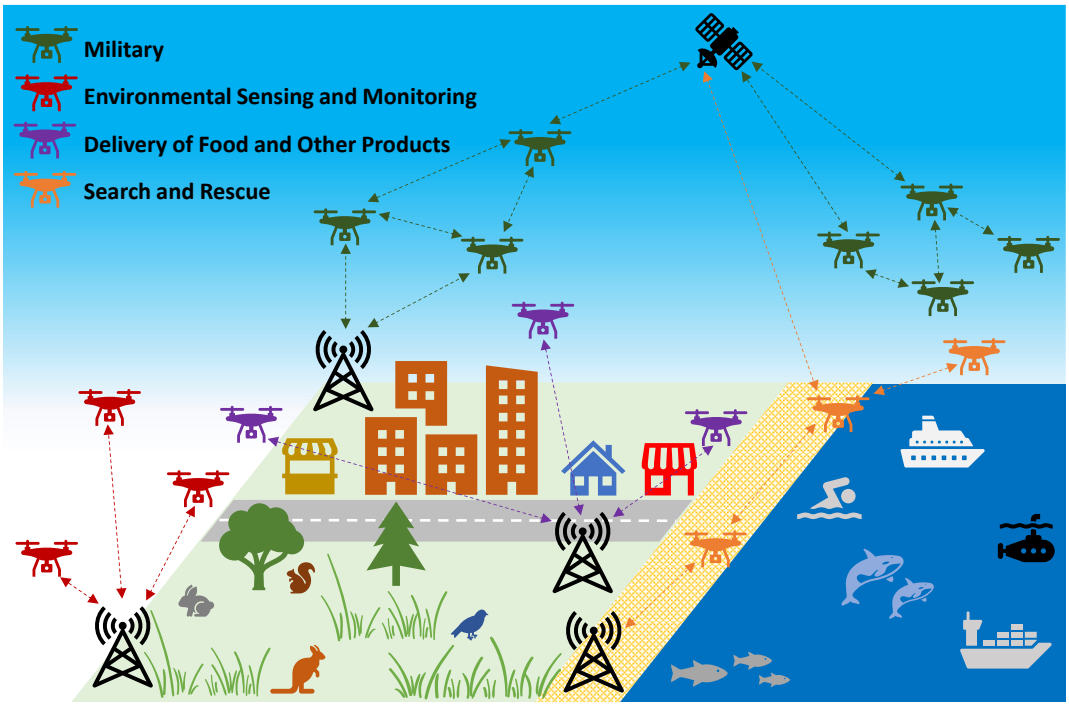

Figure 4.1: Ubiquitous UAV Deployments for Various Applications

The arrangement of UAVs to form a swarm has been increasingly highlighted as an operating model of great potential for various applications. For example, Gao et al. [17] and Zhang et al. [18] discussed the use of UAV swarms for search and rescue operations, while Liu et al. [19] discussed the use of UAV swarms for air quality index monitoring. Although the deployment of UAV swarms can bring about immense advantages from the aspects of resource allocation, control and cooperation, such a deployment model can also concurrently introduce additional security risks associated

with malicious use [122]. For instance, there could be a greater potential for attacks involving identity falsification, one of which is the Sybil attack.

Sybil attack is well-known to be one of the security threats to the IoT. It refers to the situation when a malicious node falsely claims to have numerous identities [99] [123]. There are several incentives for a node to act in such a way; in the context of FANETs, examples are such as to allow it to illegitimately acquire more weight in a voting system and to create an illusion of traffic congestion in a particular area [51] [52]. Countermeasures for Sybil attack include prevention, detection and mitigation. Prevention refers to the inhibition of the attack from occurring at all. Detection refers to the identification of security breach, the identification of attack type, as well as the initiation of relevant mitigation solutions. Finally, mitigation refers to the alleviation of resulting outcomes of the attack [124].

More recently, the use of machine learning (ML) has increasingly been leveraged to address various challenges, including IoT security. Machine learning does this by intelligently choosing the actions to be taken in response to a given situation based on knowledge that the system has learned. Well-known examples of applications are such as computer vision, bio-informatics, fraud/malware detection, authentication and speech recognition [76].

As will be discussed further in Sections 4.2 and 4.3, there exist numerous studies in the literature that discuss Sybil attack detection methods for wireless ad hoc networks, wireless sensor networks and vehicular ad hoc networks (VANETs). However, this is not the case for FANETs, which would have had relatively fewer Sybil attack threats due to the lower expectation of having high node density presented in an area; but the more recent increase in UAV usage is changing all that. Adapting one of the numerous existing non-FANET Sybil attack detection methods is also deemed to require significant effort, as those schemes were not designed to suit nodes with complex three-dimensional mobility. These facts motivated the development of a

novel approach for Sybil attack detection to fill this gap, which should be lightweight, highly secure, and able to detect smart malicious nodes with power control capability. Machine learning has been identified as a tool with high potential to aid in the delivery of such identified features.

In this chapter, a new intelligent Sybil attack detection approach for FANETs-based IoFT is proposed. The proposed novel approach employs range-based location verification using physical layer characteristics of the radio signals emitted from the UAVs as detected by two ground nodes. This is done by utilising a supervised machine learning approach and experimenting with several different classifiers available in the Weka [125] workbench platform. The learning is carried out on two features of the radio signals, namely, the received signal strength difference (RSSD) and the time difference of arrival (TDoA).

The technical contributions of this chapter are summarised as follows:

- To fill a knowledge gap in the literature relating to Sybil attack detection in FANETs-based IoFT which is still quite deficient in general.
- To achieve Sybil attack detection in FANETs-based IoFT using intrinsically generated physical layer data of radio signals emitted from the UAVs. Advantages associated with this are such as less susceptibility to attacks involving information spoofing and not requiring additional communications overheads.
- To achieve Sybil attack detection in FANETs-based IoFT, where both classic malicious nodes with fixed power and smart malicious nodes with power control capability may be presented.
- To investigate and demonstrate the use of machine learning in carrying out Sybil attack classification determination based on two attributes, namely, the

RSSD and TDoA ratios of two different radio signals, obtained using only two monitoring nodes.

The remainder of this chapter is organised as follows. Section 4.2 reviews existing related works on position localisation using physical layer data, Sybil attack detection in IoFT and machine learning systems. Section 4.3 then discusses the motivations and deduces the contributions of this study. Section 4.4 outlines the details of the proposed scheme. Section 4.5 describes the simulation environment, including all the key simulation parameters. The results and evaluation of the simulation are then discussed in Section 4.6. Finally, conclusion is given in Section 4.7.

4.2. Related Works

4.2.1. Position Localisation Using Physical Layer Data

Many existing positioning systems are known to function using measurements of physical layer features of the radio signal. Very commonly used features include received signal strength (RSS), angle of arrival (AoA), time of arrival (ToA) and time difference of arrival (TDoA). Classical usage of these measurements involves a two-step process, as briefly described in the subsequent paragraphs. To assist with visualisation, Figure 4.2 has been included to add a simplified graphical overview of the described mechanisms, as relevant to FANETs. Interested readers can also refer to more comprehensive publications, such as Dardari et al. [126] and Munoz et al. [127] for more details, including mathematical descriptions. Furthermore, additional details around RSS and TDoA, as relevant to the proposed scheme in this chapter, are described in Section 4.4.

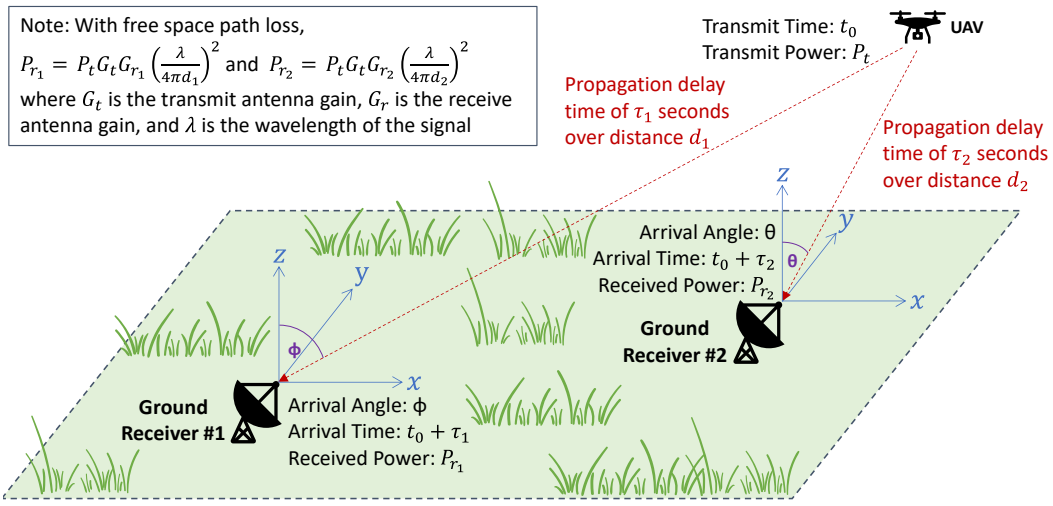

Figure 4.2: Physical Layer Position Localisation Mechanisms – FANETs

In the first of the two-step process, the position-related signal parameters of interest are measured. Out of the four features, RSS is known to be the most easily obtainable because it is simply a measurement of the received power, which can easily be done in any system without the need for time synchronisation. On the other hand, ToA and TDoA require some sort of time synchronisation. In essence, ToA is a measurement of signal propagation delay; therefore, time synchronisation between the receivers and the transmitter would be required. Similarly, TDoA is a measurement of signals propagation delay difference between the receivers; therefore, time synchronisation between the receivers of interest would be required. On the other hand, AoA is known to perhaps be the least favourable feature, as it requires characterisation of the direction of signal propagation; consequently, the use of AoA may dictate the need for costly specialised hardware, such as the use of antenna arrays. Additionally, AoA position estimation performance also degrades as the distance between transmitter and receiver increases [126] [127].

The second step is the application of position estimation techniques based on the parameters obtained in the first step. This can be achieved by using techniques such as lateration and angulation. The use of multiple types of position-related parameters can also be combined to form hybrid methods [126]. One constraint of this step is that

more than two receiver nodes are generally required for accurate positioning. For example, as outlined by Li et al. [128], according to the principles of trilateration, if ToA or TDoA are used, three receiver nodes would be required for two-dimensional position estimation. More relevant to FANETs is the fact that for three-dimensional position estimation, four receivers would be required.

4.2.2. Sybil Attack Detection in IoFT

There are quite a number of published articles that outline different Sybil attack detection methods that are applicable to slightly different IoT domains; several recent survey papers summarise these into their associated categories. Recent surveys on Sybil attack detection in wireless ad hoc networks and wireless sensor networks can be found in Arshad et al. [124], Vasudeva and Sood [129], and Singh [130]. There are also several recent survey papers on Sybil attack detection in VANETs, including Shobana and Arockia [131], Zhang et al. [132], Velayudhan and Anitha [133], and Hammi et al. [134].

Existing Sybil attack detection approaches found in the literature include the use of location verification, network behaviour monitoring, resource testing, trust systems and cryptography. As mentioned in Section 4.1, the scheme proposed in this chapter focuses on the range-based location verification approach. To elaborate further, the location verification approach is classified into range-free and range-based methods. In the range-free methods, high accuracy location is calculated based on data supplied through external means, such as Global Positioning System (GPS), radar or other localisation schemes. The range-based methods, however, generally can work simply by using data obtainable from the physical layer characteristics of the radio signals being sent and received [130]. There are several reasons why methods that use intrinsically generated physical layer data to detect Sybil attack might be more preferable than others. For instance, the use of intrinsically generated physical layer

data also brings about a security advantage over methods that use extrinsic data, in that such use would be less susceptible to spoofing attacks. Furthermore, unlike many other methods in other detection approaches, authentication would not be required; consequently, misidentification due to potentially stolen credentials would be less of a risk. Cryptography, which is a widely used technique for authentication, also consumes a lot of energy [52]. Accordingly, since UAVs operate on limited energy, for some applications, it may be desirable to cut down on their cryptographic usage. Nevertheless, there may be other advantages associated with the other Sybil attack detection methods; therefore, in some situations, it may be desirable to combine the advantages associated with different schemes by using two or more detection mechanisms on a complementary basis.

Schemes that use physical layer characteristics of the radio signals do exist in the literature. These schemes use features such as RSS, AoA, ToA and TDoA for their location verification determination. However, none of these schemes are designed for mobile nodes that possess aviation characteristics like UAVs in FANETs. In fact, apart from those designed for VANETs, most schemes only cater for static nodes. Additionally, most schemes also do not cater for the situation in which malicious nodes can adjust their transmit power to fool detectors while carrying out Sybil attacks. Furthermore, the ways in which some of these schemes operate impose various other undesirable constraints. For example, schemes such as Kabbur and Kumar [135] and Yuan et al. [136] use RSS indication values obtained through triangulation, requiring at least three monitoring nodes to be used [124] [134]. Other examples include schemes like Lv et al. [137], Abbas et al. [138] and Angappan et al. [139], which require the use of additional localisation information such as those obtainable through neighbours of the suspicious nodes [124]; consequently, unlike schemes that purely and directly use intrinsically generated physical layer data, these schemes may be more susceptible to attacks involving information spoofing.

When looking more specifically at Sybil attack detection for FANETs, there are currently no survey papers that discuss this topic. Nevertheless, a limited number of existing research works can be found in this area, including de Melo et al. [140], Sun et al. [51] and Walia et al. [141], details of which are summarised in the following paragraphs. Note that none of these schemes operate on pure use of physical layer characteristics of the radio signals.

In de Melo et al. [140], an identity and location validation scheme called UAVouch is proposed to detect malicious UAVs that do not follow expected trajectories, including the potential scenario where a Sybil attack is being carried out. The idea is for this scheme to supplement the authentication mechanism by requesting position validation from neighbouring nodes inside a cell and by using a position plausibility/classifier model to detect movement inconsistencies. The scheme is reported to have an average position falsification attack detection accuracy of above 85%.

In Sun et al. [51], a Bayesian Nash equilibrium game theory-based intrusion detection scheme is proposed, which can detect Sybil attacks among other attack types. The game is between the intrusion detection nodes and the attacking nodes, with each side strategising to maximise their profits. The scheme works by studying the past behaviour of UAV nodes and determining the deployment of intrusion detection nodes to achieve optimisation by minimising the overhead while achieving a high detection rate. Specific details on the Sybil attack detection mechanism and the associated detection accuracy rate are not provided due to not being the focus of the paper.

In Walia et al. [141], a mutual authentication technique to detect Sybil attack in FANETs is proposed. The scheme works by having each node checking its neighbouring nodes for identification. If nodes with the same identification but with different neighbours are found, they are marked as intruder nodes. Each intruder node is then

monitored more closely and if found to change its identity then it would get identified as malicious. In terms of performance, the paper reports high throughput, low overhead and low packet loss; however, it does not mention the overall Sybil attack detection accuracy rate.

4.2.3. Machine Learning for Sybil Attack Detection in IoFT

A typical machine learning system has three layers: 1) input; 2) feature extraction and processing; and 3) output. The input layer takes in pre-processed data, which is then passed onto the feature extraction and processing layer where the data patterns get extracted; basically, this is where the training of a machine learning system takes place. Several classifiers exist in this layer, each of which defines a different methodology for data pattern extraction; well-known ones are such as Support Vector Machines (SVM), Principal Component Analysis (PCA), and Hidden Markov Model (HMM). Finally, the output layer produces the prediction results of the task, such as classification for discrete outputs (class labels) and regression for continuous numeric outputs [76] [80].

Machine learning methods can commonly be grouped into *supervised*, *unsupervised*, *semi-supervised* or *reinforcement learning* approaches. Interested readers can refer to survey papers such as Jamalipour and Murali [142], Sarker et al. [143], Farooq et al. [144], Hussain et al. [76], Al-Garadi et al. [82], and Wang et al. [145] for more information on these machine learning approaches and on the use of machine learning in IoT security in general. Of most relevant to the study in this chapter is the supervised learning approach, where a class label is assigned to identify each data entry in the training set. Learning then takes place based on this known identification and the other input features parameters. Subsequently, the learned system can be deployed on other datasets to make predictions regarding the correct class label associated with each entry.

4.3. Motivations and Contributions

As can be seen from previous sections, there is currently a gap for a Sybil attack detection mechanism that can achieve highly accurate detection of mobile Sybil nodes in FANETs-based IoFT. This is especially true if the scheme can detect Sybil nodes with power control capability. In addition, the use of physical layer features was identified as potentially being very useful for Sybil attack detection applications in FANETs. The pure use of intrinsically generated physical layer data to carry out detection also minimises potential problems such as the risk of data spoofing. A potential approach might be to try and adapt existing methods developed for wireless ad-hoc networks, wireless sensor networks or VANETs to cater for FANETs; however, significant extensions would be required and there is no guarantee that such solutions will work well. As an alternative, it might be worth investigating a new innovative scheme. These reasons, together with the existence of machine learning systems as a potential solution to intelligently detect Sybil attack instances in FANETs-based IoFT, motivate the invention of a state-of-the-art Sybil attack detection scheme proposed in this chapter.

From the perspective of selecting the most appropriate physical layer features to use, the use of RSS and/or TDoA features makes the most sense. The use of ToA is undesirable because it requires synchronisation with the transmitter, which would be impractical to implement. Similarly, the use of AoA feature would also be impractical unless antenna arrays are already required for other reasons. From the amount of monitoring nodes perspective, it would also be desirable to minimise these while still maintaining a highly accurate detection functionality.

As discussed above, it was identified that there is a potential for machine learning to be used to aid the construction of a Sybil attack detection scheme. More specifically, it is known that RSS and TDoA features capture some location information. A machine

learning system can be developed to learn certain characteristics associated with RSS and TDoA values confirmed as belonging to Sybil attack events, in preparation for it to identify similar malicious instances in the future. More importantly, the learning can be performed without the system requiring to know the exact underlying mechanisms, such as mathematical operations. Because machine learning can easily learn from both features concurrently, a hypothesis can be formed that a minimal number of two monitoring nodes may already be sufficient for accurate Sybil attack detection functionality. It is important to note that the exact formats of attributes to be fed into the machine learning system need to be refined to suit the intended application, which is Sybil attack detection in this case. This process is a bit of an artwork, and for this study, it resulted in two attributes, namely, the RSSD and TDoA ratios of two different radio signals, more details of which can be found in Section 4.4.

As will be further demonstrated in later sections, the scheme proposed in this chapter, incorporating an artificial intelligence mechanism, has been designed with the intention of filling the gap for Sybil attack detection in the FANETs environment. The proposed scheme addresses all of the above-mentioned design criteria and does not require any additional communications overheads. With the use of only two monitoring nodes at fixed locations while still able to achieve a high detection accuracy of above 91% on average, it supports the hypothesis that such a minimal number of nodes may already be sufficient when assisted by a machine learning mechanism. To provide further illustration, Table 4.1 summarises the contribution of the proposed scheme compared with the existing Sybil attack detection approaches described by Singh [130].

Table 4.1: Comparison of the Proposed Scheme to Existing Sybil Attack Detection Approaches

Sybil Attack Detection Approach	Approach Description	Typical Advantages	Typical Disadvantages
Range-based location verification	Use data obtainable from the physical layer characteristics of radio signals being sent and received.	• Low in cost since device already has physical layer characteristics of communicating radio signals by default.	• Accuracy may be reduced by rapid changes in node position. • Difficulties in detecting nodes that can manipulate signal strength. • Accuracy may be reduced by interference, multipath fading, shadowing, etc.
Range-free location verification	Location calculated from external data (e.g., GPS, radar, etc).	• Can provide high accuracy distance calculation.	• External data means more susceptibility to data spoofing attacks (compared with range-based location verification).
Network behaviour monitoring	Based on nodes features and behaviour in the network.	• Allow features and behaviour in the network to be used for accurate detection of malicious nodes.	• Malicious nodes with specific knowledge can escape detection. • Specialised tools required for data collection and analysis.
Resource testing	Node challenged to provide knowledge about specific resources (usually physical fingerprinting or energy).	• Allow uniqueness in resources of each node to be used for verification.	• Extensive power consumption. • Genuine nodes with resource problems due to other reasons may be falsely classified as malicious.

Sybil Attack Detection Approach	Approach Description	Typical Advantages	Typical Disadvantages
Trust systems	Trust value obtainable from trusted devices or trusted neighbours must be maintained by each node to remain in the network.	• Allow periodic evaluation which can be done in centralised or decentralised manner.	• Inability to detect malicious node already dominating trust determination process.
Cryptography	Authenticate nodes and communicate securely using public/private keys. Use watermarking to guarantee valid data.	• Can also offer protection against various other attack types.	• High memory, computing, and communications overhead for resource constraint devices. • High costs associated with key management.
Proposed scheme	Range-based location verification using physical layer characteristics of the radio signals, namely, RSSD and TDoA, paired with supervised machine learning.	• Low in cost since device already has physical layer characteristics and only require two monitoring nodes. • Less susceptible to attacks on upper layers, such as data spoofing, stolen credential, etc. • Designed to work with mobile nodes in FANETs-based IoFT. • High detection performance even with malicious nodes that can manipulate signal strength. • Potential to extend to detect other attack types and/or utilise unsupervised machine learning approach.	• Prior to deployment, some further performance studies may still be required, for example on: 1. effects of interference and structural blockages; and 2. networks with high node density.

4.4. System Architecture

In this section, the architecture of the proposed scheme is discussed. As depicted in Figure 4.3, this study looks at a situation where a number of UAVs fly within a given area to carry out certain operations. While doing so, the UAVs communicate with each other and/or with ground stations. Some members of the nodes have malicious purposes and would attempt to carry out Sybil attacks by falsely identifying themselves as other entities. Two monitoring nodes are placed on the ground at fixed locations within the operational area in an attempt to detect Sybil UAV nodes.

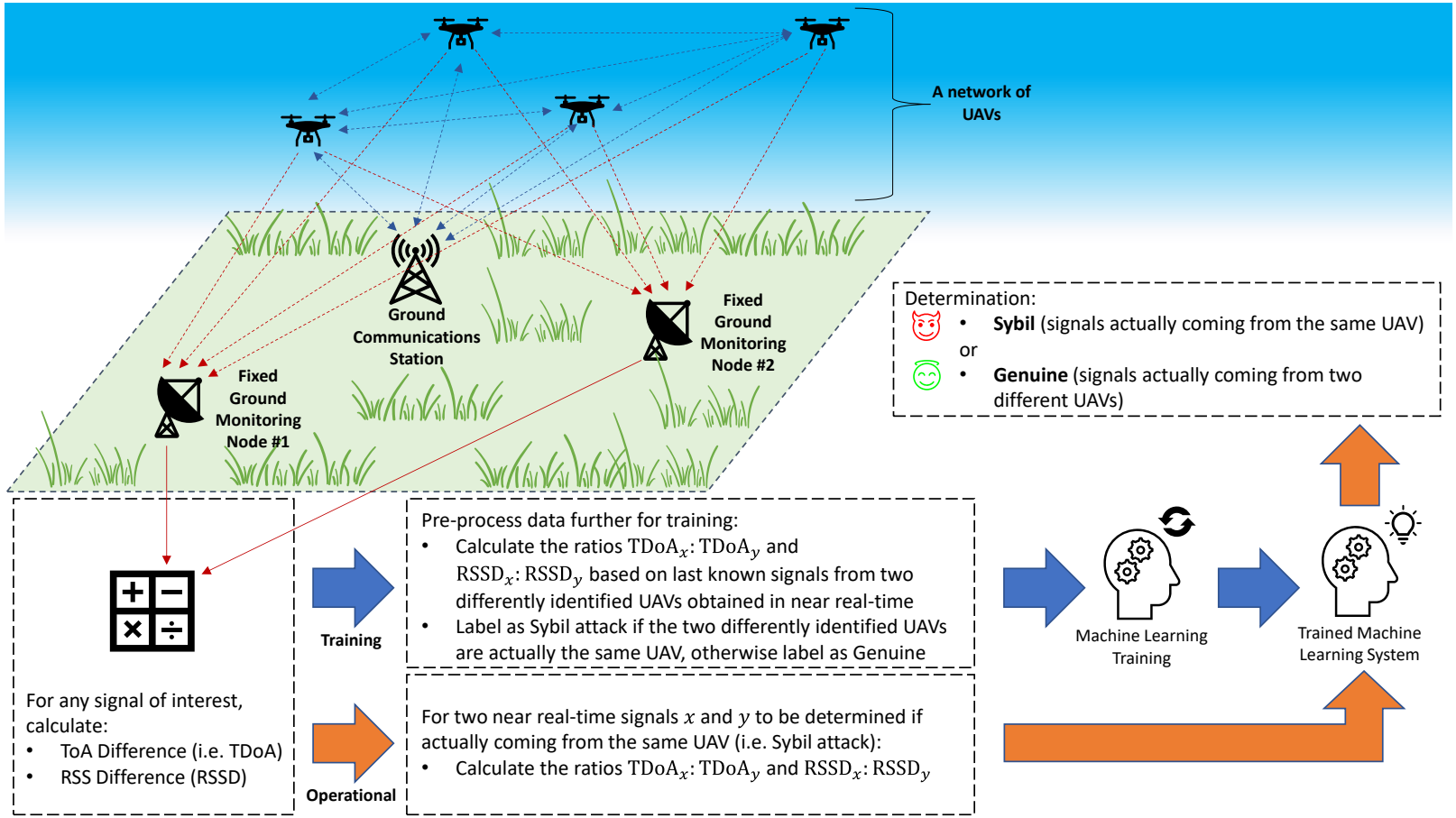

Figure 4.3: Architecture of the Proposed Scheme

The focus of the architecture is on the use of machine learning system to detect Sybil attack instances. The supervised machine learning approach was identified as the most suitable approach in this study due to the nature of the problem being addressed. This is because there are simply two known distinct outcome classes, which are whether or not a Sybil attack event is taking place. Furthermore, the use of supervised learning is also favourable from the performance assessment perspective, as training and test datasets with correctly labelled class events can be generated in a straightforward manner through the simulation of UAV networks.

As outlined in Section 4.2.3, a typical machine learning system has three layers: 1) input; 2) feature extraction and processing; and 3) output. In this architecture, the focus is mostly on the input layer, more specifically, the derivation of data attributes to be fed into the machine learning system. Feature extraction and processing activities, which result in the determination of classification output, are mostly performed by the machine learning system based on specific algorithms. There exist numerous well-researched supervised machine learning algorithms which can potentially be used with the proposed architecture, as long as they support two numerical attributes (i.e., RSSD and TDoA ratios) and a class attribute (i.e., Sybil attack instance or not). Some of these algorithms have been selected for the simulations carried out in this study, the details of which can be found in Sections 4.5.3 and 4.5.4.

Before proceeding further, it is important to note that the proposed scheme has been designed with the intention of being flexible for use with a range of UAV mobility patterns, density levels, transmit power levels and signal emission rates; however, the exact limitations are outside the scope of this study. Another point to note is that this study was conducted based on the assumption that the free space path loss propagation model holds true. Furthermore, it is also assumed that signals from other UAVs and other systems in the surrounding area are coordinated in such a way that results in negligible interference effects on the functionality of the system, such as through the use of orthogonal frequency-division multiplexing.

Regarding how the machine learning attributes were designed, as discussed in Sections 4.2 and 4.3, the literature review carried out suggests that RSS and TDoA physical layer features contain location information most suitable for the application scenario in this study. However, this is not the end of the process as the exact formats of machine learning input attributes that would allow for a high classification success rate still need to be derived. Based on the assumptions given in the previous paragraph, the received power level is assumed to follow the free space path loss model developed by Friis [146] as

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d}\right)^2 \quad (1)$$

where P_t is the transmitted signal power, G_t is the transmit antenna gain, G_r is the receive antenna gain, λ is the wavelength of the signal, and d is the distance between transmitter and receiver. As for the signal propagation time taken between the transmitter and the receiver, such delay can be represented by the equation

$$\tau = \frac{d}{c} \quad (2)$$

where the constant $c \approx 3 \times 10^8$ m/s can be used for the speed of light [126].

The proposition given in this Sybil attack detection problem is that there are to be only two monitoring nodes, and the system is to detect if two signals identified as transmitted from different UAVs in near real-time are actually likely coming from the same location (i.e., the same UAV). Therefore, it is necessary to ensure that the attributes are designed to capture maximal information to enable the machine learning classifier to recognise such underlying pattern differences. As the use of machine learning is an experimental science, the process of determining the precise formats of the attributes used in this study requires some creativity and preliminary experiments to verify their effectiveness. Because of the proposition to use two monitoring nodes, the use of TDoA measured at these different monitoring nodes already makes sense. The next step is to represent this characteristic as a numeric

value that captures the relationship between the two signals. The numeric value needs to somehow encompass such a relationship pattern to make it distinguishable if the value is likely coming from two signals belonging to the same UAV. Consequently, the TDoA ratio between the two signals was selected as one of the attributes. Similarly, the process also resulted in the RSSD ratio of the two signals as the other attribute, where the RSSD of any signal of interest is the difference between the RSS values of the signal measured at the two different monitoring nodes. The precise details of these two attributes are captured in Algorithm 4.1 and Algorithm 4.3. Since the differences in RSS and ToA values received at the two monitoring nodes play important roles in the characterisation of the two attributes, it is worth noting that the two locations should be sufficiently far apart to enhance effectiveness.

There are two different phases in this scheme: the *training* and *operational* phases. Detailed descriptions are elaborated in the following subsections.

4.4.1. Training Phase

In the training phase, UAVs would be deployed and carry out radio communications in a controlled manner. A given number of these UAVs would be programmed to act maliciously and execute Sybil attacks by falsely using multiple identities. Signals from each UAV would be sampled by the monitoring nodes every certain interval for a given number of times until the end of the training period. The two monitoring nodes would detect and collect the RSS and ToA of each signal being sampled. Subsequently and in accordance with Algorithm 4.1, the corresponding RSSD and TDoA of each signal sampled would be calculated based on the variations in RSS and ToA received at the two different nodes. This is followed by the execution of Algorithm 4.2, which also calls Algorithm 4.3, to calculate the RSSD and TDoA ratios between the signal being sampled and all other latest signals sampled from every other UAV with a differently

declared identity. The data generated by Algorithm 4.2 for all collected signals would then be collectively fed to the machine learning classifier as training data.

As previously discussed, the Sybil attack characterisation being performed is carried out through the discovering of patterns within the RSSD and TDoA ratios from two different signals that would have been emitted from somewhat nearby physical positions during the UAV's movement in the air. Consequently, it is also important to note that any signals that were emitted in the past beyond a certain near real-time threshold need to be excluded, as the positions of the UAVs emitting those would potentially have already changed quite significantly. This threshold is represented by the near real-time limit t_{thres} in Algorithm 4.2.

4.4.2. Operational Phase

In the operational phase, UAVs would be deployed and carry out communications using either genuine or fake identities. In this phase, the two monitoring nodes would collect RSS and ToA data of radio signals and use Algorithm 4.1 to calculate RSSD and TDoA similarly to the training phase; however, the difference is that the true identity of each signal's emitter is not known. To predict whether any two near real-time signals collected and identified as coming from different UAVs are actually coming from the same UAV (i.e., a Sybil attack event), Algorithm 4.3 is executed and the resulting values of RSSD and TDoA ratios are passed on to the trained machine learning classifier for determination.

There are various ways in which the operational phase detection mechanism can be deployed. As an example, detection can be performed on all pairs of signals detected by the monitoring nodes and the machine learning classification results are passed on to the upper layers for appropriate risk-based decisions subjecting to other relevant information available. Alternatively, perhaps more efficiently, an individual request

can be made by a mechanism in one of the upper layers to perform a check on any particular signals suspicious of being from a Sybil node.

Algorithm 4.1: Calculate TDoA and RSSD for a given signal x

Input: $\mathrm{ToA}_{x\#1}$: ToA of radio signal x obtained by fixed ground monitoring station #1,
$\mathrm{RSS}_{x\#1}$: RSS of radio signal x obtained by fixed ground monitoring station #1,
$\mathrm{ToA}_{x\#2}$: ToA of radio signal x obtained by fixed ground monitoring station #2,
$\mathrm{RSS}_{x\#2}$: RSS of radio signal x obtained by fixed ground monitoring station #2

begin

1 Calculate $\mathrm{TDoA}_x = \mathrm{ToA}_{x\#1} - \mathrm{ToA}_{x\#2}$

2 Calculate $\mathrm{RSSD}_x = \mathrm{RSS}_{x\#1} - \mathrm{RSS}_{x\#2}$

end

Algorithm 4.2: Execute Algorithm 4.3 on a given signal x and all the latest signals collected from each of the other differently identified UAVs within a given near real-time limit t_{thres}

Input: Current time t,
Near real-time threshold t_{thres},
UAV_{ID}: the identity of UAV claimed to have emitted signal x

```
1  begin
2      for each collected signal claimed to have emitted from a UAV other than UAV_ID do
3          if the signal y being checked is the latest emitted from such claimed UAV identity at current time t and the signal y was not sampled prior to t − t_thres do
4              Execute Algorithm 4.3 with signal x and signal y as the two signal inputs
5              if the real identity of emitters of signal x and signal y are actually the same UAV do
6                  Mark corresponding entry as "Sybil" class
7              else
8                  Mark corresponding entry as "Genuine" class
9              end if
10         end if
11     end for
12 end
```

Algorithm 4.3: Calculate ratios $\text{TDoA}_x : \text{TDoA}_y$ and $\text{RSSD}_x : \text{RSSD}_y$ for two given signals x and y

Input: TDoA_x: Latest known near real-time TDoA of signal x obtained from Algorithm 4.1,
RSSD_x: Latest known near real-time RSSD of signal x obtained from Algorithm 4.1,
TDoA_y: Latest known near real-time TDoA of signal y obtained from Algorithm 4.1,
RSSD_y: Latest known near real-time RSSD of signal y obtained from Algorithm 4.1

```
begin
1      Calculate TDoA_x : TDoA_y = TDoA_x ÷ TDoA_y
2      Calculate RSSD_x : RSSD_y = RSSD_x ÷ RSSD_y
end
```

4.5. Simulation Environment

A simulation of the proposed scheme was set up on a desktop computer with an Intel i7 2.90 GHz processor, 32 GB of random access memory (RAM), and Windows 10 Enterprise operating system. The simulation can be divided into three stages: 1) simulation of a network of flying and communicating UAVs; 2) data pre-processing prior to machine learning classification; and 3) machine learning classification. The network simulator OMNeT++ [115] (Version 5.7) was used in conjunction with the INET framework [147] (Version 4.2.9) for the first stage, the output of which is a log file containing all communication records. Subsequently, for the second stage, a Python script was written and applied to the log file. This was performed to extract all relevant data, execute relevant algorithms described in Section 4.4, and arrange the collated data to a format readable by the machine learning classifier used in the next stage. Finally, in the third stage, machine learning classification was carried out using the previously prepared training and test data. The tool used for the third stage was the Weka workbench platform (Version 3.8.5). Details of the three stages and further information on the Weka workbench platform are described in the following subsections.

4.5.1. Stage 1: Simulation of UAVs

In this stage, a network of flying and communicating UAVs was simulated in OMNeT++ using INET's "MassMobility" model. The UAVs movement model was based on INET's "3D Mobility" showcase [148], in which each UAV node moves in a three-dimensional space. To summarise, the UAV nodes moved at a speed randomly selected from a uniform distribution range between 10 and 20 m/s. Each node also turns at a random uniform distribution angle range between $-10°$ and $10°$ around a random elevation angle of the same uniformly distributed angle range. The positioning of the UAVs was

configured to update every 1 s. In terms of the UAVs flying space, this was defined as a square of dimensions 1000 × 1000 m. As for the elevation, range was restricted to be between 5 to 150 m to better reflect a more realistic permitted flying height for UAVs. On the ground, three fixed nodes were added: 1) the ground communications station at coordinates (250, 400); 2) the first monitoring node at coordinates (250, 250); and 3) the second monitoring node at coordinates (750, 750).

On the communications side, INET's "AckingWirelessInterface" wireless network interface module was used together with "ApskScalarRadio" hypothetical radios and the "ApskScalarRadioMedium" radio model which uses free space path loss by default [149]. A transmission frequency of 2 GHz was specified for use with this radio model. Antenna gains were not defined, which means that an isotropic antenna with a gain of 1 (0 dB) was used for each radio [150].

Machine learning training data was simulated based on a network of 100 UAV nodes, 80 of which were genuine in that they only used their true identities to identify themselves in communications. Each genuine node transmitted one UDP packet to the ground communications station every 1 s period. The initial transmission time was different for each node, but ranged between simulation times $t = 1$ and $t = 2$ s. The other 20 UAV nodes were Sybil nodes, each of which used two different identities, namely, "A" and "B", to identify itself. Each identity transmitted one UDP packet to the ground communications station every 1 s period. Similar to genuine nodes, the initial transmission time for Sybil nodes was different for each identity, but ranged between simulation times $t = 1$ and $t = 2$ s. Note that relating back to the near real-time limit t_{thres} described in Section 4.4, the limit used here can be considered as not exceeding 1 s. The transmission period of 1 s can also be viewed either literally as each UAV identity communicated once a second, or perhaps more realistically, that each UAV identity communicated numerous times a second but only one of those got sampled.

In terms of the transmission power, the assumption was that all UAV nodes are supposed to be operating at a power level that is not too high, in order to preserve their limited onboard battery energy. At the same time, the transmit power needs to be high enough to achieve reliable radio transmission in various environments and distances. Therefore, all genuine UAV nodes were defined to transmit at a power level of 100 mW, which is also assumed to be the maximum transmit power level. Conversely, Sybil nodes had the ability to adjust their transmission power down to a smaller level in an attempt to fool more traditional Sybil attack detectors.

The training data was generated for two scenarios: 1) where each Sybil node operates at a fixed transmit power level of 100 mW; and 2) where each Sybil node operates at a fixed transmit power level of 100 mW for Identity A but at a range of power levels from 100 mW down to as low as 0.001 mW for Identity B. More specifically, power levels assigned to different Identity B UAV nodes are 100 mW, 75 mW, 50 mW, 25 mW, 10 mW, 0.1 mW and 0.001 mW. The training data simulation for each scenario was carried out for a duration of 50 simulated seconds using “seed-set” value of “0”. Such a timing duration was chosen to achieve a balance of having sufficient training data samples while minimising actual simulation execution time.

For the generation of test cases, two main different transmit power scenarios were used, similar to what were used for the training data. Likewise, an execution duration of 50 simulated seconds was also used. Nevertheless, more diverse test cases were generated; for instance, the tests include some power levels presented in the training data as well as some power levels not presented in the training data but still within the 0.001 mW to 100 mW range. For each test case, the evaluation was done on more diverse “seed-set” values, being from “1” through to “5”. Furthermore, supplementary test cases were generated for a new UAV network composition consisting of 98 genuine nodes and 2 Sybil nodes, also using various power levels within the same range and “seed-set” values of “1” through to “5”. Note that the designation “Gx80Sx20” will be used to refer to the network composition comprising

80 genuine nodes and 20 Sybil nodes. Similarly, the designation "Gx98Sx2" will be used to refer to the network composition comprising 98 genuine nodes and 2 Sybil nodes.

4.5.2. Stage 2: Data Pre-Processing Prior to Machine Learning Classification

In this stage, for both the training and test data, a Python script was written to extract all relevant data from the output log file generated by OMNeT++ and arrange the data into an "ARFF" dataset format readable by Weka. Each dataset had three attributes: 1) $\mathrm{TDoA}_x : \mathrm{TDoA}_y$ ratio; 2) $\mathrm{RSSD}_x : \mathrm{RSSD}_y$ ratio; and 3) class label of either Sybil or Genuine. The generation of these attributes using Algorithm 4.2 is described in detail in Section 4.4.

Because the simulated UAV networks consisted of a substantially higher number of genuine nodes than Sybil nodes, the generated datasets contained substantially more entries of the Genuine class. This means that the machine learning classifier would learn more characteristics of Genuine class data than Sybil class data, and thus would be more susceptible to overfitting the data to the characteristics of the Genuine class nodes. To mitigate this issue, a decision was made to also create a trimmed down version of the training data which randomly skips some entries of the Genuine class so that there are roughly equal entries for the Genuine and Sybil classes overall. Some quick experiments were performed and confirmed that using the untrimmed version for training resulted in the classifier having a much poorer performance in detecting Sybil class entries. As an example, Table 4.2 illustrates the OneR classification results when using the trimmed and untrimmed training datasets for Scenario 2 described in Section 4.5.1 evaluated against the trimmed and untrimmed versions of one of the Gx80Sx20 test datasets. Note that the details of how this table was populated can be referred to in Section 4.5.3. Unsurprisingly, the use of untrimmed training data led to

very high true positive detection rates of Genuine class entries but very low true positive detection rates of Sybil class entries. Although such use led to a very high average overall accuracy percentage when evaluated with the untrimmed test dataset, this was only so because there were significantly more instances of Genuine class data. As can be seen, when using such untrimmed training data evaluated with the trimmed test dataset, the average overall accuracy percentage was very low. Similar results were also obtained with the use of different test datasets and classifiers. Consequently, a decision was made to use the trimmed version of the data for training. For testing, although it may be more realistic to use the untrimmed data, a decision was made to also experiment with the trimmed data for the Gx80Sx20 composition in order to observe the machine learning classification performance more thoroughly.

Table 4.2: OneR Classification Results – Training Dataset Evaluated With a Gx80Sx20 Test Dataset – Trimmed vs. Untrimmed

Training Dataset	Overall Correct Classification Percentage	True Positive Sybil Instances	False Negative Sybil Instances	True Positive Genuine Instances	False Negative Genuine Instances	True Positive Sybil Percentage	True Positive Genuine Percentage
	Evaluated with untrimmed version of test dataset:						
Trimmed	91.07%	1837	83	622412	61108	95.68%	91.06%
Untrimmed	99.71%	66	1854	683380	140	3.44%	99.98%
	Evaluated with trimmed version of test dataset:						
Trimmed	94.16%	1837	83	1775	141	95.68%	92.64%
Untrimmed	51.67%	66	1854	1916	0	3.44%	100%

4.5.3. Stage 3: Machine Learning Classification

In this stage, evaluation is carried out on the training and test data in Weka. The Weka platform comes included with a collection of classifiers of different algorithm types. Furthermore, additional classifiers are also available as optional downloadable

packages. In this study, preliminary experiments were carried out with most, if not all, of the classifiers that support the problem scenario, in an attempt to shortlist a few high-performing ones. This process then narrowed down to the four chosen algorithms, namely, J48, Classification via Regression, OneR, and JRip. Note that these well-researched algorithms are of three different types, more details of which can be found in Section 4.5.4 for interested readers. These diverse algorithms were then used in carrying out the full experiments to observe the robustness of the scheme.

The following output results were captured for evaluation: 1) the accuracy of correctly classified instances overall; 2) the number of Sybil class entries correctly identified as Sybil class (i.e., "true positive Sybil" or equivalently "true negative Genuine"); 3) the number of Sybil class entries incorrectly identified as Genuine class (i.e., "false negative Sybil" or equivalently "false positive Genuine"); 4) the number of Genuine class entries correctly identified as Genuine class (i.e., "true positive Genuine" or equivalently "true negative Sybil"); and 5) the number of Genuine class entries incorrectly identified as Sybil class (i.e., "false negative Genuine" or equivalently "false positive Sybil").

Note that the second and third outputs can be used to calculate the percentage of true positive Sybil entries detection. Similarly, the fourth and fifth outputs can be used to calculate the percentage of true positive Genuine entries detection. Another point to note is that the second and fourth outputs can be added together to obtain the overall correct classification instances. Likewise, the third and fifth outputs can be added together to obtain the overall incorrect classification instances.

4.5.4. Weka Workbench Platform

The Weka workbench platform is a popular open-source software for machine learning [151] [152]. Weka comes with a collection of classifiers, where this study focuses on the following four: 1) J48; 2) Classification via Regression; 3) OneR; and 4)

JRip. These four classifiers are based on three different algorithm types: 1) decision tree; 2) metalearning; and 3) rules. These different classifiers are briefly described below.

The J48 classifier is a decision tree type algorithm. Decision trees define the sequences of decisions to be made together with the resulting recommendation. Each node in a decision tree evaluates a specific attribute until a leaf node is reached, which is where the classification decision is made. The J48 classifier is a derivation of a straightforward divide-and-conquer algorithm called "C4.5" [153] which needed to be extended in order to cater for real-world problems [154].

The Classification via Regression classifier is a metalearning type algorithm. Metalearning algorithms take classifiers and make them into more powerful learners or change them for other applications [154]. In the case of the Classification via Regression classifier, it performs classification on discrete classes using regression methods which would otherwise only be suitable for continuous classes. Note that the M5P decision tree classifier [155], which is the default option, was used in the experiments carried out in this study.

The OneR and JRip classifiers are rules type algorithms. Rules-based classifiers are popular alternatives to decision trees. Rules can be much more consolidated than decision trees, especially when it is possible to have a default rule covering cases not defined by other rules. Another reason for rules popularity is that new rules can be added to existing ones without disrupting the other rules already in place [154].

The OneR classifier, which is also called "1R" or "1-rule", is Weka's implementation of Holte [156]. It works based on a set of rules applied to just one attribute by creating a different set of rules for each attribute and choosing the best one based on the resulting error rates. It is described as a simple and efficient method that can still produce effective rules that can often achieve surprisingly high accuracy. An explanation for such a phenomenon is that often the pattern underlying any real-

world data is quite fundamental that even only just one attribute of the data is adequate for performing accurate predictions [154].

The JRip classifier is Weka's implementation of the Repeated Incremental Pruning to Produce Error Reduction (RIPPER) rule learner [157]. It is based on the idea of using incremental reduced-error pruning by Fürnkranz and Widmer [158] for quick and effective rule inference [154].

4.6. Simulation Results and Evaluation

This section looks at machine learning classification results obtained from the simulation exercises described in Section 4.5, where the results for all test datasets were obtained from five simulation runs using the five different "seed-set" values. To summarise, a high correct classification accuracy of above 91% on average was achieved across all four selected machine learning algorithms, even in scenarios with smart malicious nodes operating at power levels not directly trained. Such a high performance reflects the suitability of the design choices made for the proposed architecture, especially the selection of the two machine learning attributes, namely, the RSSD and TDoA ratios of two different signals. Additionally, the results also reflect the robustness of the proposed architecture in upholding high performance when different machine learning classifiers are used.

The following two subsections discuss the results in more detail. Note that this study uses three criteria for the evaluation metrics (refer Section 4.5.3): 1) the accuracy of correctly classified instances overall ("correct classification accuracy"); 2) the percentage of true positive Sybil entries detection ("true positive Sybil rate"); and 3) the percentage of true positive Genuine entries detection ("true positive Genuine rate").

4.6.1. Sybil Nodes With Fixed Transmit Power Level

This subsection examines the performance of Scenario 1, the results of which are shown in Figure 4.4. This is a simpler scenario in which Sybil nodes can only transmit at a fixed power level of 100 mW. It can be seen that the correct classification accuracies exceed 96% for all classifiers except for OneR which performs slightly worse in this scenario but still exceeds 91%. Similar results can also be observed when looking more specifically at true positive Sybil and true positive Genuine rates. Another observation about OneR is that it also performs worst in terms of its equitability in distinguishing Sybil and Genuine class entries, with the gaps between the true positive Sybil and true positive Genuine detection rates being the largest among the three classifiers.

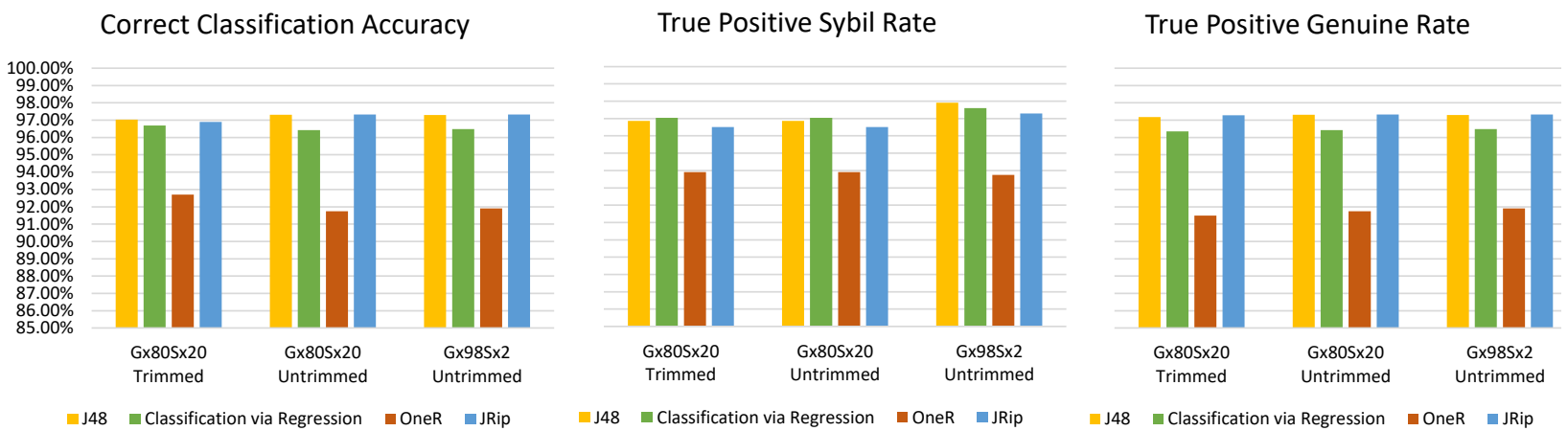

Figure 4.4: ML Classification Results – Sybil Nodes With Fixed Transmit Power Level

4.6.2. Sybil Nodes With Variable Transmit Power Level

This subsection considers Scenario 2, which represents the more complex cases where Sybil nodes can vary their transmit power level. For these cases, a training dataset containing Sybil nodes with seven different transmit power levels was used to train the classifiers. Testing was conducted using diverse datasets with various transmit power levels, some of which were included in the training dataset and some of which were not.

4.6.2.1. Average Results

To characterise the results more generally, the average results obtained from the use of all test datasets for each of the two different node compositions, as shown in Figure 4.5, are examined. When compared with the fixed power results shown in Figure 4.4, it can be seen that the correct classification accuracies of the four classifiers decrease by a few percent, but all still exceed 91%. Likewise, the true positive Sybil and true positive Genuine rates also decrease slightly, with the results for true positive Sybil appearing to be slightly higher than that of true positive Genuine for all classifiers; however, the gaps are smallest for the JRip classifier, indicating that it is the most equitable one in distinguishing Sybil and Genuine class entries. Interestingly, unlike the results for the fixed power scenario, the performance of the OneR classifier is now more similar to that of the other three classifiers. This is perhaps not too surprising because as outlined in Section 4.5.4, OneR only uses one attribute to create rules, and so the performance in some situations would be worse than the other classifiers that use all attributes available.

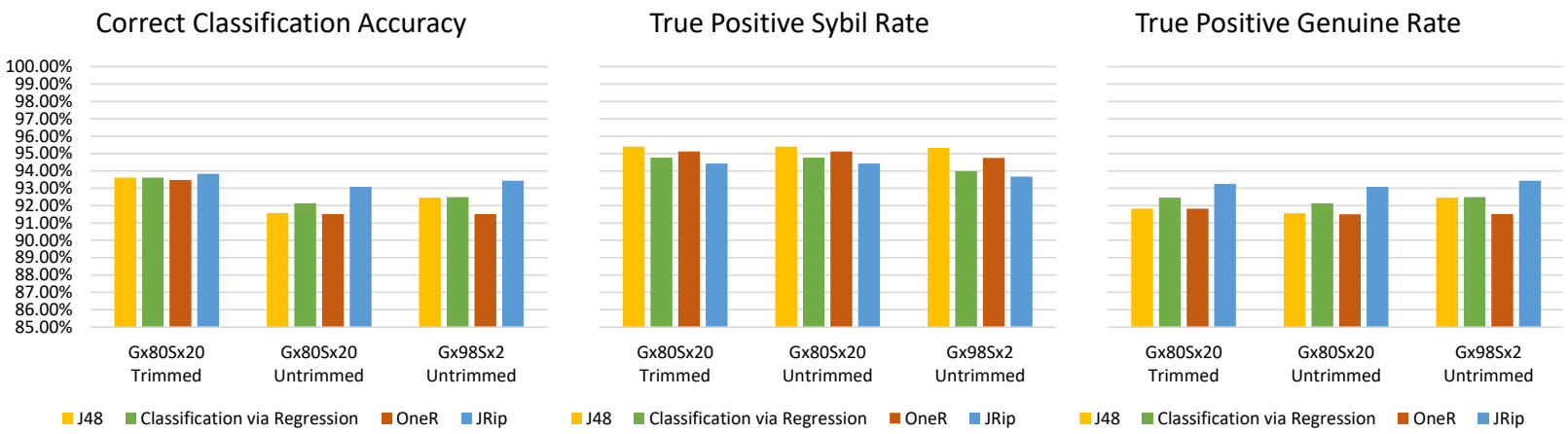

Figure 4.5: ML Classification Results – Average Results for Sybil Nodes With Variable Transmit Power Level

4.6.2.2. More Detailed Samples of Results

This subsection looks more closely at the performance differences between some sample test cases where the transmit power levels of Sybil nodes were included in the training dataset versus those that were not.

Firstly, consider an example situation where the Sybil nodes can only transmit at a fixed power level that is already included in the training dataset. Figure 4.6 illustrates the classification execution results of one such example situation, where each Sybil node transmits at a power level of 50 mW for its Identity B. When comparing this with the average results shown in Figure 4.5, it can be seen that the results are fairly consistent with one another. The correct classification accuracies of all four classifiers exceed 91%. Similarly, the true positive Sybil rates are only slightly higher than the true positive Genuine rates for all classifiers.

Next, consider a situation where the transmit power level of each Sybil node's Identity B has not been included in the training dataset. An example situation is illustrated in Figure 4.7, which captures the results of a diversified test case where there are five different transmit power levels of Identity B used among the Sybil nodes, namely, 40 mW, 3 mW, 0.6 mW, 0.03 mW and 0.007 mW. Note that such a situation was only created for the Gx80Sx20 node composition, and not for the Gx98Sx2 node composition, because the small number of Sybil nodes in the latter case would not be effective in demonstrating the intended diversification. In terms of the classification results comparison, it can be seen that the results are also in line with the average results captured in Figure 4.5, where the correct classification accuracies for all classifiers exceed 91%.

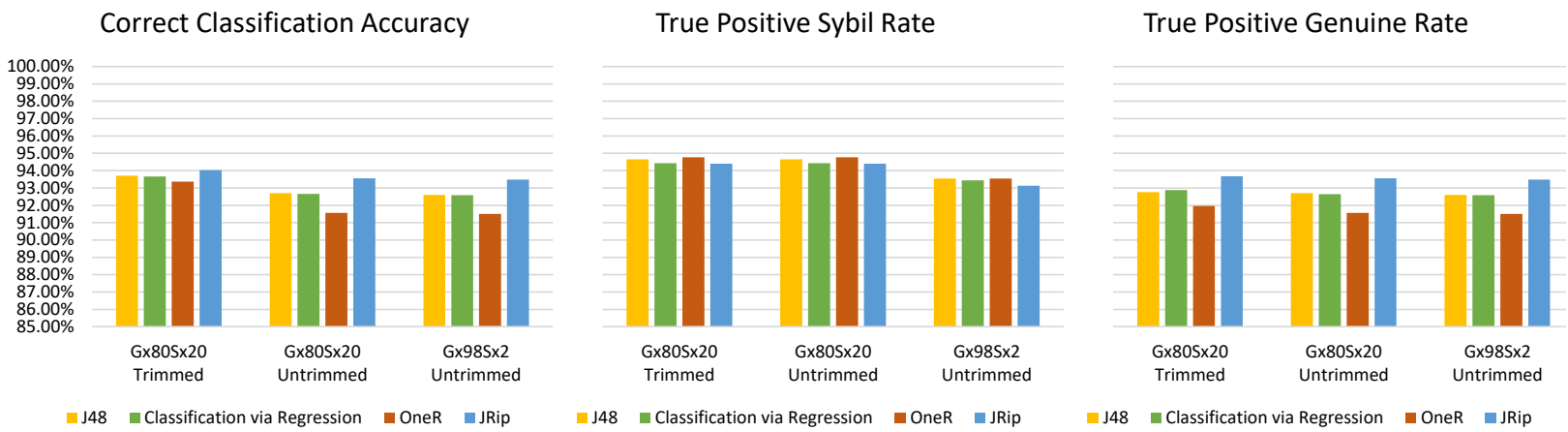

Figure 4.6: ML Classification Results – Sybil Nodes With Trained Identity B Transmit Power Level (50 mW)

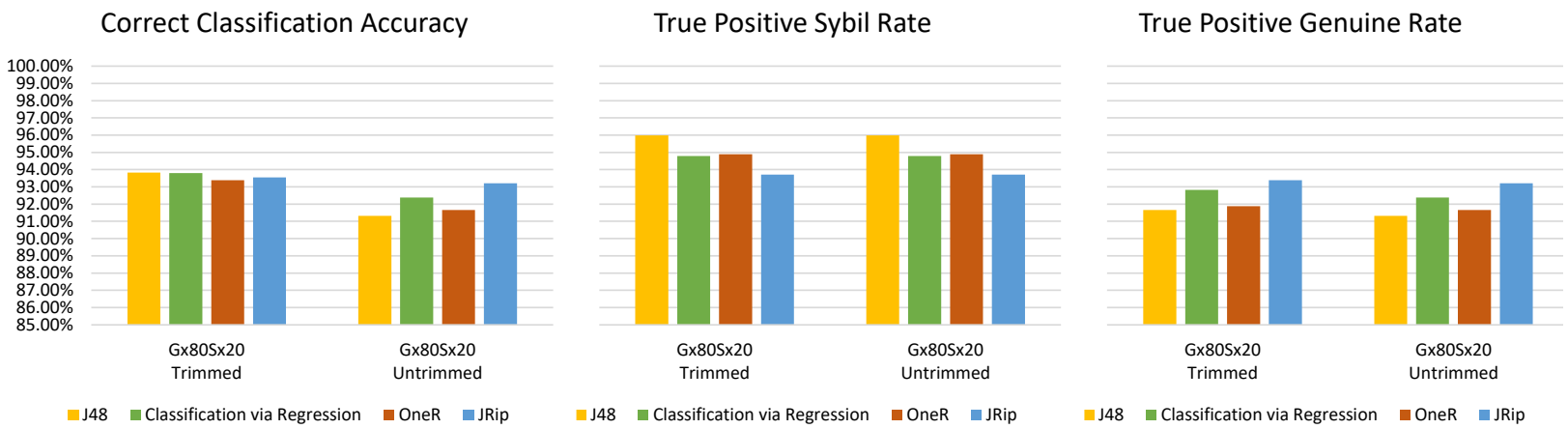

Figure 4.7: ML Classification Results – Sybil Nodes With Untrained Identity B Transmit Power Levels (Mixture of 40 mW, 3 mW, 0.6 mW, 0.03 mW and 0.007 mW)

4.6.3. Future Works

The experimental results in this study were obtained from simulations carried out based on several assumptions which may not necessarily hold true in all situations. Therefore, more considerations are required prior to the actual deployment of the proposed scheme and may necessitate further experiments and adaptations, as appropriate. In addition, the proposed scheme may also be extendable to provide improvements and additional functionalities.

4.6.3.1. Additional Considerations Prior to Deployment

Examples of issues that may require additional consideration prior to deployment are as follows. Firstly, this study was carried out based on the assumption that the free space path loss propagation model holds true and that signals from other systems in the surrounding area are coordinated in such a way that results in negligible interference effects on the functionality of the system, such as through the use of orthogonal frequency-division multiplexing. Consequently, further assessments would need to be done on the effects of interference and structural blockages applicable at the physical location the system is planned to be deployed in.

This study also uses a specific UAV mobility model taken from an INET framework's showcase, which defines how different UAVs move around in a range of random speeds and directions. Simulations were also performed on specific flying space dimensions and node density levels. Furthermore, the simulations carried out used only one near real-time threshold value which is $t_{\text{thres}} = 1$ s. In practice, depending on the application, it is possible that nodes may be required to fly in a higher density environment. They may also be required to use different mobility patterns or signal

emission rates. Therefore, further studies are needed on these aspects prior to deployment, as appropriate.

The UAVs transmit power levels used in this study range from the maximum value of 100 mW down to the minimum value of 0.001 mW. Although these values led to great simulation results, further investigations would need to be done to confirm the performance based on the expected minimum and maximum UAVs transmit power levels applicable to the deployment scenario.

4.6.3.2. Use of Alternative Machine Learning Attributes

The RSSD and TDoA ratios of two different radio signals were selected as the attributes used for machine learning in this study following the hypothesis that they capture significant location information regarding any given UAV node at a particular point in time when used together. From the simulation results, the use of these two attributes was found to be quite effective in detecting Sybil attacks. Nevertheless, more studies can be carried out in the future to investigate whether the performance of the scheme can be improved even further if additional and/or different attributes are used, including those derived from other physical layer features, especially if such other features are easily obtainable in the intended deployment scenario.

4.6.3.3. Extension to Support Unsupervised Machine Learning and Other Attack Types

In reality, there may be situations in which datasets for machine learning training are not easily obtainable. In such situations, the use of supervised machine learning may not be ideal. As a potential solution, unsupervised machine learning, which uses input datasets without class labels to independently extract useful information and patterns [142], may need to be considered as an extension of the scheme. Likewise,

considerations should be given to extending the scheme to cater for other attack types in FANETs, a good starting point of which might be those that also involve location verification.

4.6.3.4. Adaptation to Support Other Application Scenarios

Notwithstanding the fact that the proposed scheme was designed for and experimented in the FANETs environment, the approach may also function well in other application scenarios, either as is or with some modifications. As an example, in the case of VANETs, the mobility patterns where vehicles of certain height travel on known roads can be considered two-dimensional, which is more restrictive than the three-dimensional mobility in FANETs. However, there are similarities that may enable the mechanisms underlying the proposed scheme to also function well in such an environment. Additionally, research on VANETs is also more mature and thus trusted infrastructures exist, such as roadside units (RSUs), which may be advantageous for the adaptation of the proposed scheme (e.g., the RSUs can potentially be used as ground monitoring nodes).

4.7. Conclusion

This chapter proposes a supervised machine learning approach to intelligently detect Sybil attacks for FANETs-based IoFT. Simulation results revealed that the proposed scheme can achieve a high correct classification accuracy of above 91% on average, even for smart malicious nodes with power control capability operating at power levels not directly trained. Correspondingly, this means that the proposed scheme has a low false classification rate of less than 9% on average. Additionally, because of the use of only intrinsically generated physical layer data, the proposed scheme is also less susceptible to various attacks commonly carried out on the upper layers, such as data

spoofing. Furthermore, no additional communications overheads of the UAV nodes are required for the functionality of this scheme. For future works, it may be possible to extend this scheme beyond Sybil attack detection applications, for example, to address other problems in FANETs that involve location verification. In addition, extensions and adaptations to support unsupervised machine learning and other application scenarios can also be investigated.

Chapter 5 Inter-Pulse Interval for Frequency Hopping Sequence Determination

Human body interface and control systems (HBICS) refer to the information exchange between devices inside, on, and within the proximity of a human body. Research in HBICS security is a very significant area because many types of HBICS devices are considered to be safety critical, especially those used for medical applications. Denial-of-service (DoS) through the launch of wireless communications link jamming is known to be a major attack type in HBICS. The use of inter-pulse interval (IPI) biometrics to address various HBICS authentication and encryption security challenges is known to be quite well-researched; however, existing approaches cannot be adopted for use in frequency hopping applications due to the fundamental difference in how frequency hopping operates. This gap triggers the exploration of a new approach in this chapter, to potentially enable IPI to be used to add another layer of protection to the traditional pseudorandom frequency hopping system.

5.1. Introduction

The communications among human body interface and control systems (HBICS), also known as wireless body area networks (WBAN), refers to the information exchange between devices inside, on, and within the proximity of a human body. In other words, a substantial focus of this is on human wearable and implantable devices. Applications of HBICS can range from medical to non-medical. Examples of medical applications include the monitoring and control of health conditions, such as fatigue, asthma, diabetes, cardiovascular diseases, cancer detection, and so on. Examples of non-

medical applications include entertainment, non-medical emergency management, and security management [20].

Security is one of the aspects of HBICS that often gets discussed in the literature. This is not surprising for such devices that are used for medical, emergency and security applications. One well-known case that also highlights such concern is the fact that the US Vice President Dick Cheney disabled the wireless functionality of his heart implant pacemaker due to fear of assassination through the device being hacked [26] [54] [55].

The use of biometrics in HBICS has been widely discussed in the literature as being potentially suitable for various security applications. Common types of physiological features that come into play for such applications include electrocardiogram (ECG), photoplethysmogram (PPG), fingerprint and iris [24] [26] [28] [54] [59] [61] [62]. The timing between heartbeats, also known as the inter-pulse interval (IPI), is an ECG-based mechanism, and is perhaps one of the most prominently discussed physiological biometrics. IPI has a clear benefit, in that it can be measured anywhere on the body of a person. Furthermore, another advantage of using IPI as a physiological entropy source is the fact that it has a high level of randomness [55].

Confidentiality, *integrity* and *availability* are some of the most often cited attributes in HBICS security discussions [20] [21] [26] [55] [56] [57] [58]. The use of IPI-based physiological parameters has quite extensively been proposed for authentication and encryption operations, which mitigate threats to confidentiality and integrity. However, discussions on the use of IPI-based physiological parameters to mitigate threats to availability have so far been quite limited.

The jamming of communication links to cause denial-of-service (DoS) is one type of attacks to availability. Frequency hopping can be used to counteract such an attack [24] [60]; however, if the parameters that determine the hopping pattern become compromised and somehow made known to the attacker, the attacker would be able

to easily carry out jamming attacks. This chapter proposes the use of IPI biometrics to add another layer of protection to the traditional pseudorandom frequency hopping system, which would be suitable for use in scenarios involving communications among HBICS devices implanted inside and/or worn on a human body. As part of the study, experiments were carried out based on four different algorithms, as detailed in further sections.

The remainder of this chapter is organised as follows. Section 5.2 outlines existing related works on the use of IPI for security applications and on the use of frequency hopping for anti-jamming. Section 5.3 outlines the motivations and contributions of the proposed scheme. The operational scenario of concern and the proposed algorithms are then described in Section 5.4. In Section 5.5, the simulation environment used for the experiments is discussed. This follows by Section 5.6 which discusses the simulation results, performance evaluation and security analysis. Finally, the chapter concludes in Section 5.7.

5.2. Related Works

5.2.1. Inter-Pulse Interval (IPI) for Security Applications

The idea of using IPI biometrics for HBICS security was first introduced by Poon et al. [65] in 2006 [54] [55] [62]. Its functionality is enabled through usage in conjunction with the fuzzy commitment scheme [63], which allows for errors in what is equivalent to a decryption key, to be tolerable to a certain degree [65].

The fuzzy commitment scheme [63] is known to be suitable for biometric applications owing to its support for small variabilities in physiological signals. It is also commonly adopted in other biometric authentication applications, such as the use of fingerprints as cryptographic keys [159].

Figure 5.1 illustrates a high-level conceptual overview of how biometrics are typically used in conjunction with the fuzzy commitment scheme for authentication and encryption purposes.

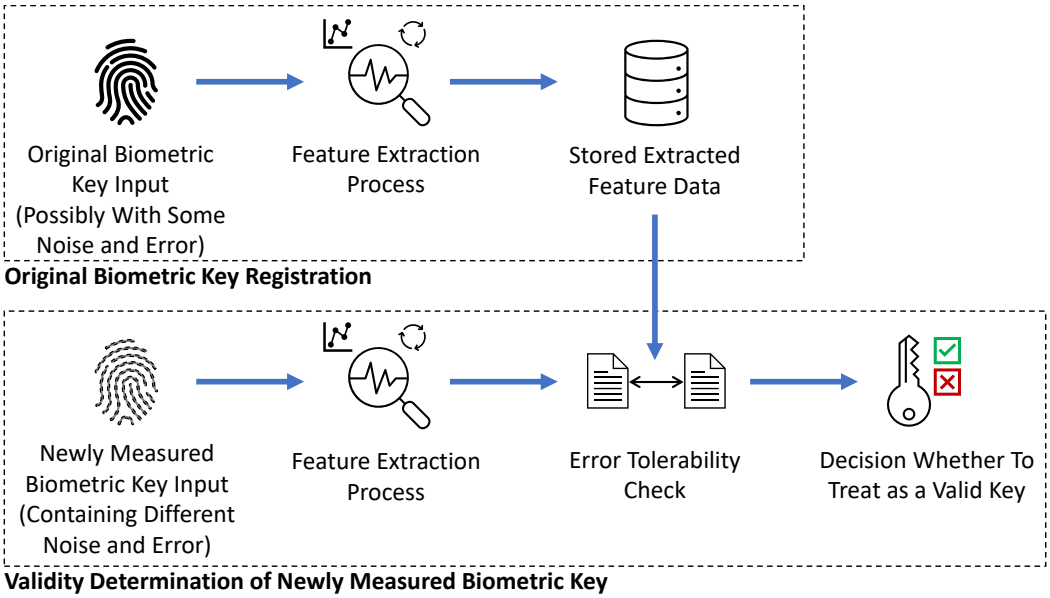

Figure 5.1: Biometrics for Authentication and Encryption – A High-Level Overview

5.2.2. Frequency Hopping for Anti-Jamming

Frequency hopping is known to be the most frequently used method to counteract frequency jamming attacks [160]. Traditionally, as illustrated in Figure 5.2, the frequency hopping pattern is determined in a pseudorandom fashion through the use of frequencies set and hop sequence commonly known to the participating time-synchronised devices [24] [60]. Consequently, this means that if the pseudorandom seed is compromised and known to the attacker, who presumably already has knowledge of other details of the system, the attacker would be able to determine which frequency devices communicate on at any given point in time. As a result, this knowledge would enable the attacker to easily carry out jamming attacks on the devices operating frequencies.

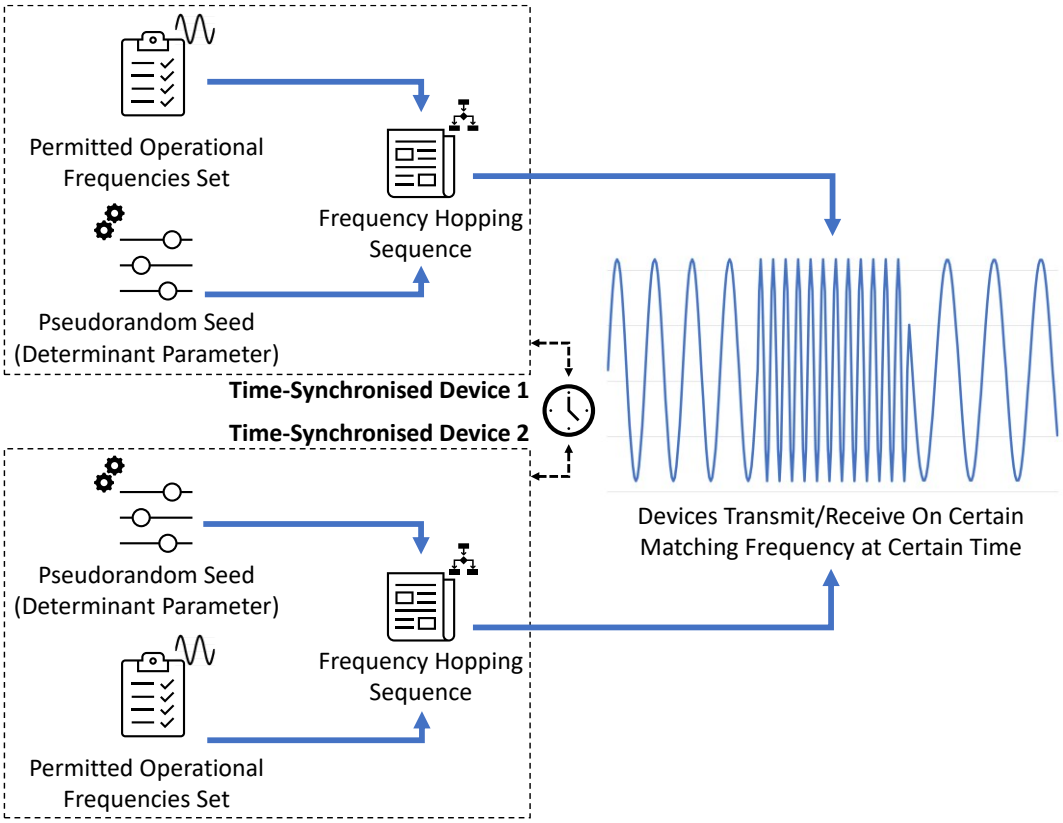

Figure 5.2: Traditional Frequency Hopping System

5.3. Motivations and Contributions

From previous sections, it can be seen that wireless communications link jamming is a security threat to HBICS, which can be mitigated through the use of frequency hopping. Due to the high level of randomness associated with IPI biometrics and the fact that IPI can be measured anywhere on the body, its potential use for frequency hopping pattern determination would provide additional security to the wearer/implantee of HBICS devices. Nevertheless, although the use of IPI biometrics for security applications in HBICS has been well-researched, existing solutions rely on validating the comparison of two sets of physiological signals that are supposed to match each other within a certain degree of error allowance. The problem is that this

type of solution, which would have worked fine with authentication and encryption applications, cannot be adopted for use in frequency hopping applications due to the fundamental difference in how frequency hopping operates. For frequency hopping applications, there is no opportunity for the frequency determinants of each of the communicating devices to be compared against each other. The frequency determinants need to be derived separately for each device with the intent to match each other with minimal errors, to ensure that the devices communicate on a matching frequency at any given point in time. This gap in suitable solutions motivates the development of a new approach which can relate IPI biometrics to frequency hopping patterns, as proposed in this chapter.

5.4. System Architecture

5.4.1. Operation Scenario

Figure 5.3 demonstrates the operation scenario of concern. In this scenario, a HBICS devices user has two or more wearable and/or implantable devices with direct body contacts. The devices need to communicate with each other for various critical tasks. As an example, these tasks may be for critical healthcare monitoring and control. An attacker with malicious intent may wish to disable the functionality of one or more implantable devices to cause deteriorating health to the user. If the devices of interest communicate on a known frequency, the attacker can achieve this by transmitting on the same channel with high enough transmit power to jam the communication link being used, causing service deterioration. As already introduced in previous sections, frequency hopping can be used to counteract such an attempt. However, in a traditional frequency hopping system, the hop pattern is determined in a pseudorandom manner. Therefore, if the attacker can somehow manage to get hold

of the relevant pseudorandom determinant parameters, jamming attacks can still be easily carried out. To mitigate such concern, the ongoing IPI signals intrinsically generated by the user, which are available to the HBICS devices through direct body contacts, can be used to contribute to the determination of frequency hopping sequence.

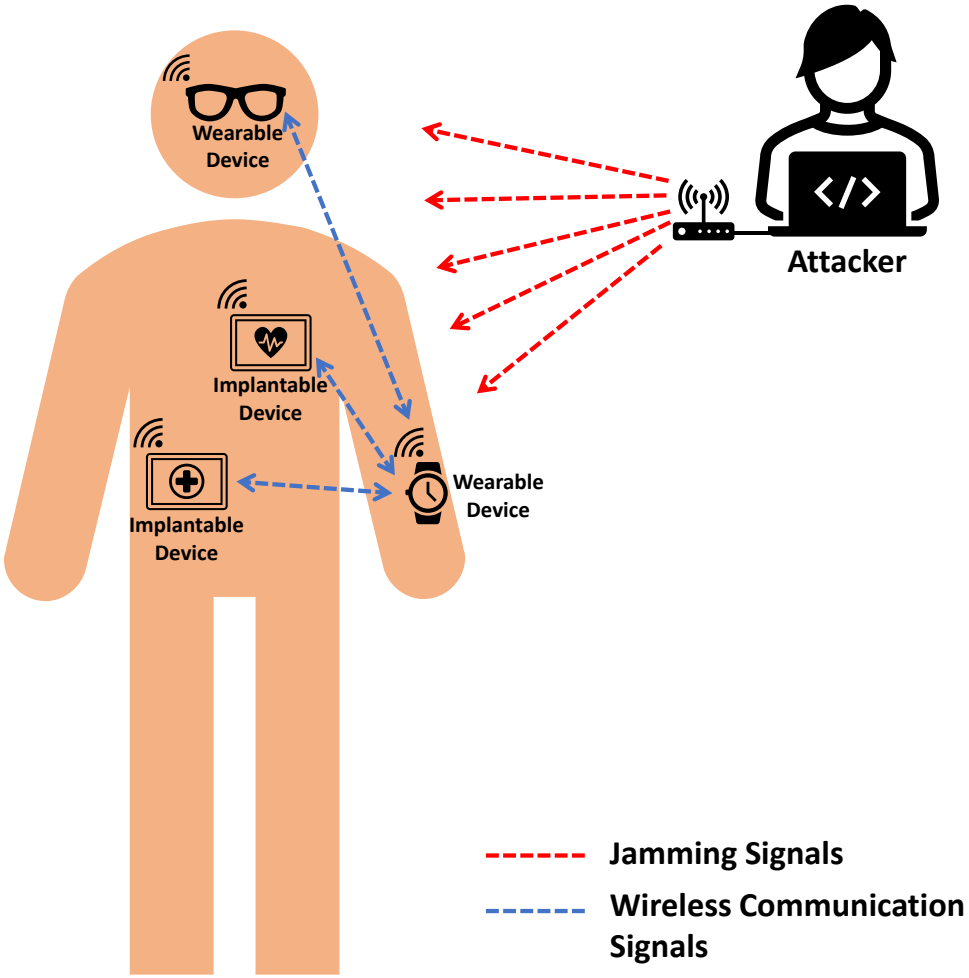

Figure 5.3: Operation Scenario of Concern

5.4.2. Proposed Algorithms

Four alternative algorithms, from Algorithm 5.1 through to Algorithm 5.4, are hereby proposed for use with IPI biometrics to add another layer of protection to the pseudorandomly determined frequency hopping pattern, as shown in Figure 5.4.

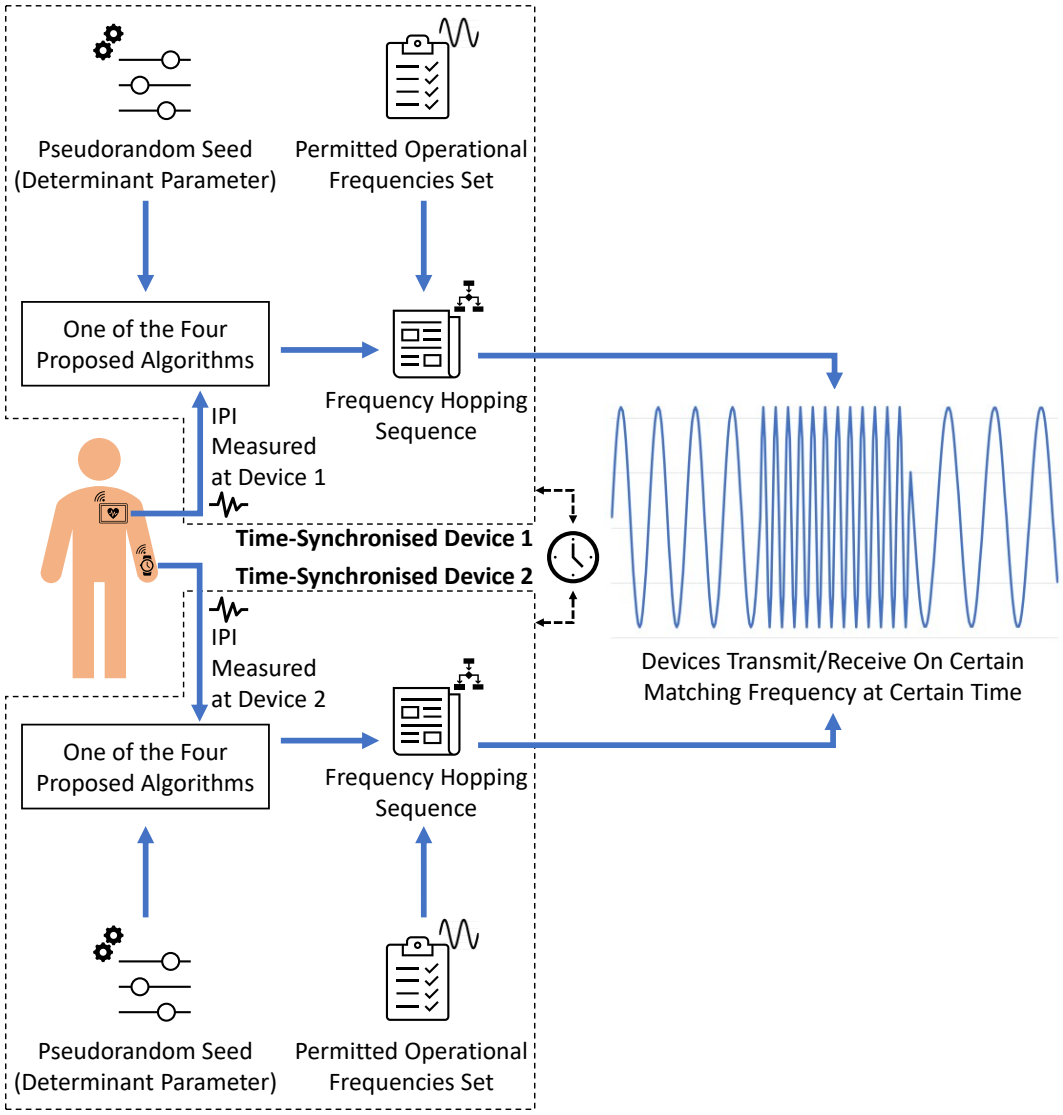

Figure 5.4: Proposed Integration of IPI and Frequency Hopping System

The proposed algorithms have been designed to work with time-synchronised HBICS devices which share a mutually known pseudorandom determinant parameter. All

four algorithms output a binary value b, given the following inputs: consecutive IPIs set P for n pulses, comparison set size s, and pseudorandom determinant parameter k. The objective of the algorithms is to output the value b that is tolerant to noise/errors caused through IPI measurements being taken at different parts of the body, which is one of the biggest barriers in allowing IPI biometrics to become functional in security applications. The value b, which gets independently calculated by each device, is to be used for determining the radio frequency to communicate on at any point in time.

Algorithm 5.1 and Algorithm 5.2 are based on the comparison of consecutively accumulated IPIs at two different points in time. Algorithm 5.3 and Algorithm 5.4 are also based on the comparison of two accumulated IPI sets, but with the element compositions pseudorandomly determined, and thus, in most cases, would not be consecutive to each other. Algorithm 5.1 and Algorithm 5.3 use summation as the basis for comparison, while Algorithm 5.2 and Algorithm 5.4 use the gradient value of the linear line of best fit.

Note that the output b obtained from one of the four alternative algorithms can be buffered and concatenated to form a larger number, subjecting to the number of frequency choices available to hop on. The periodicity of how often a new value of b is to be generated depends on the required frequency hopping speed.

Algorithm 5.1: Compare Two Subsets of Consecutively Accumulated IPIs of Size s Using Summation

Input: Set P made up of consecutive IPIs for n pulses,
Comparison set size s, where $2 \leq s \ll n$,
Comparison points determinant parameter k

begin
  Use k and the device's synchronised time source to pseudorandomly derive two indices for set P, namely i and j, where $\{i, j\} \subseteq \{s, \dots, n\}$ and $i \neq j$
  Calculate $x =$ sum of all elements of the subset $\{p_{i-s+1}, \dots, p_i\} \subseteq P$
  Calculate $y =$ sum of all elements of the subset $\{p_{j-s+1}, \dots, p_j\} \subseteq P$
  if $x \geq y$ **do**
    return $b = 1$
  else do
    return $b = 0$
  end if
end

Algorithm 5.2: Compare Two Subsets of Consecutively Accumulated IPIs of Size s Using Trendline's Gradient

Input: Set P made up of consecutive IPIs for n pulses,
Comparison set size s, where $2 \leq s \ll n$,
Comparison points determinant parameter k

begin
  Use k and the device's synchronised time source to pseudorandomly derive two indices for set P, namely i and j, where $\{i, j\} \subseteq \{s, \dots, n\}$ and $i \neq j$
  Calculate $x =$ gradient of the linear line of best fit for elements in the subset $\{p_{i-s+1}, \dots, p_i\} \subseteq P$
  Calculate $y =$ gradient of the linear line of best fit for elements in the subset $\{p_{j-s+1}, \dots, p_j\} \subseteq P$
  if $x \geq y$ **do**
    return $b = 1$
  else do
    return $b = 0$
  end if
end

Algorithm 5.3: Compare Two Pseudorandomly Determined Subsets of Accumulated IPIs of Size s Using Summation

Input: Set P made up of consecutive IPIs for n pulses,
Comparison set size s, where $2 \leq s \ll n$,
Accumulation points determinant parameter k

begin
  Use k and the device's synchronised time source to pseudorandomly derive $I = \{i_1, \dots, i_s\}$ and $J = \{j_1, \dots, j_s\}$ which are two sets of indices for P of size s
  Calculate $x =$ sum of all I indexed subset $\{p_{i_1}, \dots, p_{i_s}\} \subseteq P$
  Calculate $y =$ sum of all J indexed subset $\{p_{j_1}, \dots, p_{j_s}\} \subseteq P$
  if $x \geq y$ **do**
    return $b = 1$
  else do
    return $b = 0$
  end if
end

Algorithm 5.4: Compare Two Pseudorandomly Determined Subsets of Accumulated IPIs of Size s Using Trendline's Gradient

Input: Set P made up of consecutive IPIs for n pulses,
Comparison set size s, where $2 \leq s \ll n$,
Accumulation points determinant parameter k

begin
  Use k and the device's synchronised time source to pseudorandomly derive $I = \{i_1, \dots, i_s\}$ and $J = \{j_1, \dots, j_s\}$ which are two sets of indices for P of size s
  Calculate $x =$ gradient of the linear line of best fit for elements in the I indexed subset $\{p_{i_1}, \dots, p_{i_s}\} \subseteq P$
  Calculate $y =$ gradient of the linear line of best fit for elements in the J indexed subset $\{p_{j_1}, \dots, p_{j_s}\} \subseteq P$
  if $x \geq y$ **do**
    return $b = 1$
  else do
    return $b = 0$
  end if
end

5.5. Simulation Environment

5.5.1. Algorithms Simulation

The proposed algorithms were simulated using MATLAB R2021a. The ECG signals used were obtained from the MIMIC-III database [161] through PhysioNet [162]. Note that the MIMIC-III database contains, among other information, various vital sign data of deidentified adult patients admitted to critical care units.

In the experiments, data samples of approximately three to three and a half minutes duration each were taken from three MIMIC-III records which have at least two simultaneous ECG signal data available. The IPI values were then calculated from these ECG signals in MATLAB using the “findpeaks” function. The datasets used are shown in Table 5.1. Additionally, the first five seconds of these ECG datasets are plotted in Figure 5.5, Figure 5.6 and Figure 5.7.

In all simulations, ECG signals II and V were used to simulate separate IPI readings from two separate HBICS devices connected to the same body of a person. For Algorithm 5.1 and Algorithm 5.2, lookup tables were created to capture all possible comparison combinations corresponding to the value of comparison set size s being evaluated. The error rate can then be calculated from the total mismatched values of output b. However, for Algorithm 5.3 and Algorithm 5.4, the total number of possible comparison combinations is significantly larger; thus, full lookup tables are deemed impractical to achieve. So, instead, these algorithms were executed repeatedly until 5000 comparison instances were achieved for each value of s being evaluated, which is a good balance between the statistical results achievable and the simulation time required. The common seed value “123456” was used for MATLAB’s random number generator to simulate the pseudorandom determinant parameter k for these two algorithms.

Table 5.1: Datasets From MIMIC-III Records Used in Simulations

	Record	Signals	From Time	To Time	Duration (minutes)	Pulses
A	3100140_0011	II and V	10:03:49.407	10:06:52.999	03:03.592	250
B	3225133_0005	II and V	10:08:42.007	10:11:59.999	03:17.992	287
C	3108324_0010	II and V	06:39:18.952	06:42:36.944	03:18.944	176

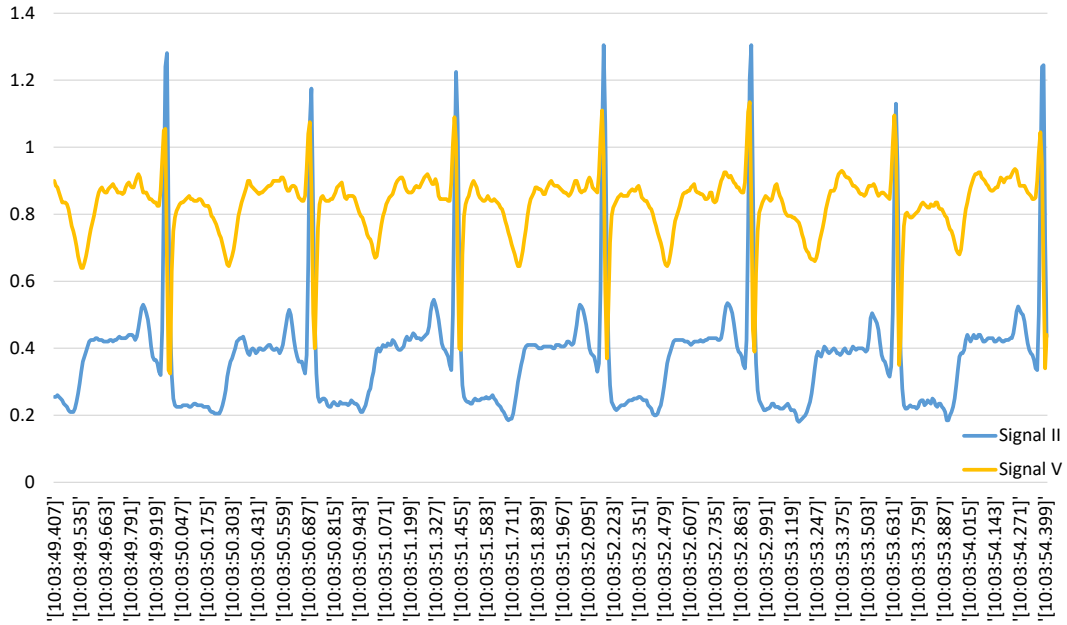

Figure 5.5: First Five Seconds of Dataset A

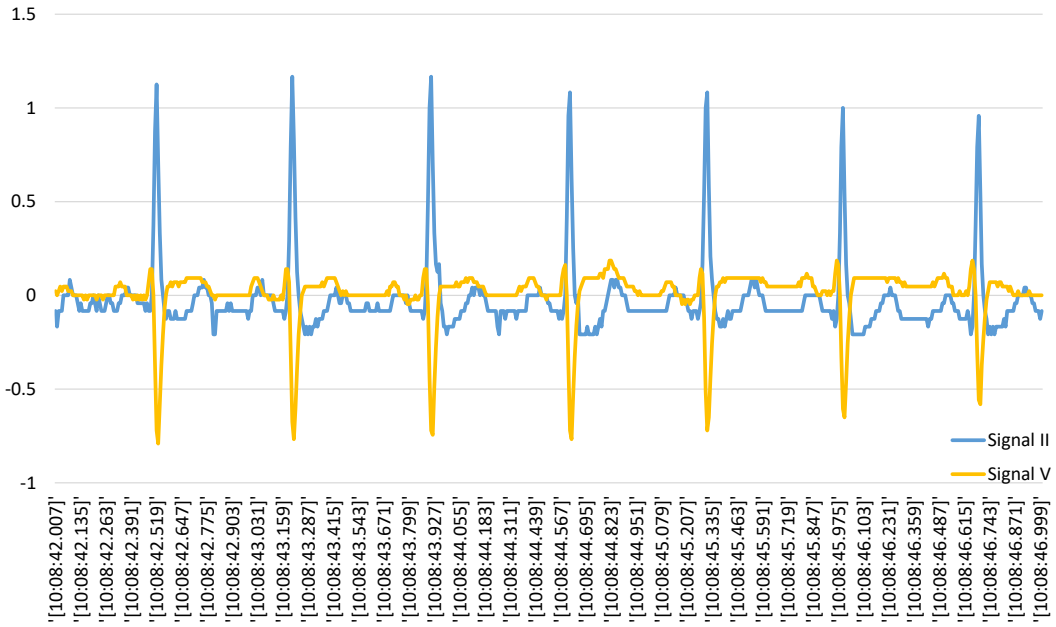

Figure 5.6: First Five Seconds of Dataset B

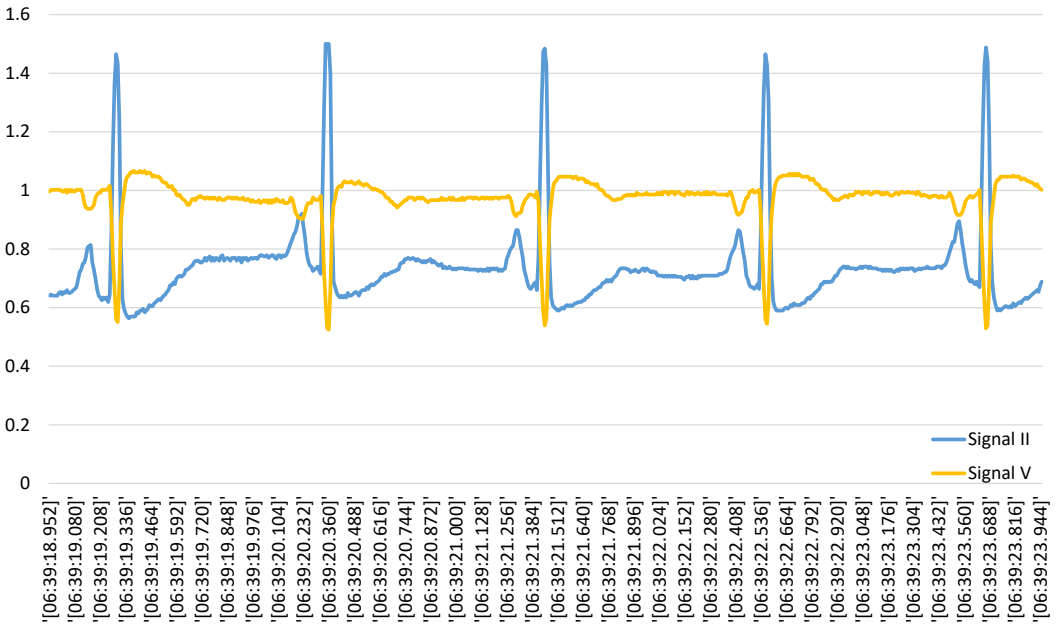

Figure 5.7: First Five Seconds of Dataset C

5.5.2. Frequency Hopping Simulation

In the frequency hopping simulation, output b was buffered and concatenated to form a 5-bit number, which was then used to determine which of the thirty-two frequency choices would be hopped on at any point in time. This simulation was done in Simulink, which is a component of MATLAB. The setup is shown in Figure 5.8, and is partially based on the Simulink frequency hopping framework presented by Fan and Tan [163]; however, unlike their framework which uses 2FSK modulation, the setup used here is based on BPSK. The simulation parameters used are listed in Table 5.2.

The frequency hopping simulation was executed based on the outputs of Algorithm 5.1 discussed in Section 5.5.1. At any given point in time, the transmitter uses the value of b obtained from ECG signal II to determine the frequency to transmit on. In contrast, the receiver uses the value of b obtained from ECG signal V to determine the frequency to receive on.

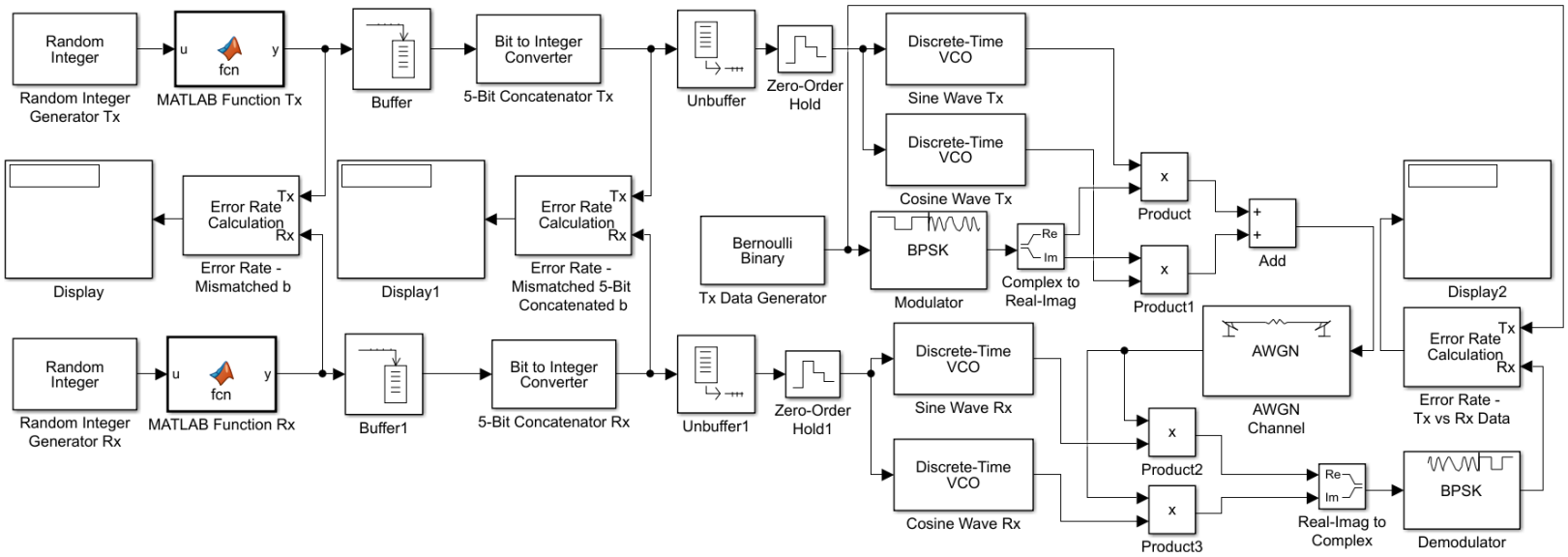

Figure 5.8: Simulink Frequency Hopping Simulation Model

Table 5.2: Simulink Frequency Hopping Simulation Parameters

Block	**Parameter**	**Value**
Random Integer Generator Tx/Rx	Set size	Total number of entries of the lookup table generated minus 1
	Initial seed	456654
	Sample time	0.004
	Samples per frame	1
MATLAB Function Tx/Rx	Input u and output y	y = value corresponding to index u in the lookup table generated
Buffer / Bit to Integer Converter	Size	5 bits / 5 bits per integer
Sine/Cosine Wave Tx/Rx	Output amplitude	1 V
	Quiescent frequency	200 Hz
	Input sensitivity	50 Hz
	Initial phase	0 rad for Sine, pi/2 rad for Cosine
	Sample time	1/4000 s
Tx Data Generator	Probability of zero	0.5
	Initial seed	61
	Sample time	0.002
	Sample per frame	1
AWGN Channel	Mode	Signal to noise ratio (SNR)
	SNR	20 dB

5.6. Simulation Results and Evaluation

5.6.1. Algorithms Simulation Results

Figure 5.9 shows the simulation results, in terms of output value b comparison mismatched percentage. The use of Algorithm 5.1 resulted in the lowest error percentage in general. This follows by Algorithm 5.2, where the error percentage reduces quite significantly for a relatively higher value of s. The use of Algorithm 5.3 and Algorithm 5.4 resulted in a significantly higher error percentage than Algorithm 5.2 in most cases, especially for datasets A and B.

Another potential issue with the proposed algorithms is the fact that the comparison $x \geq y$ would cause $b = 1$, rather than $b = 0$, to be returned when x is equal to y. This indicates that the output value may be slightly skewed towards $b = 1$. Such situations of imbalanced output values could impact security in terms of the increase in predictability. Consequently, attention needs to be paid to such potential problems in the simulation results. Figure 5.10 illustrates the percentage of instances where the output value is $b = 1$, as averaged across II and V signals. Ideally, it would be preferable to have this as close to 50% as possible. From the results, it can be seen that all data fit within the range of between 49% and slightly above 55% for a relatively lower value of s, with the range decreasing to between 49% and slightly above 51% for a relatively higher value of s.

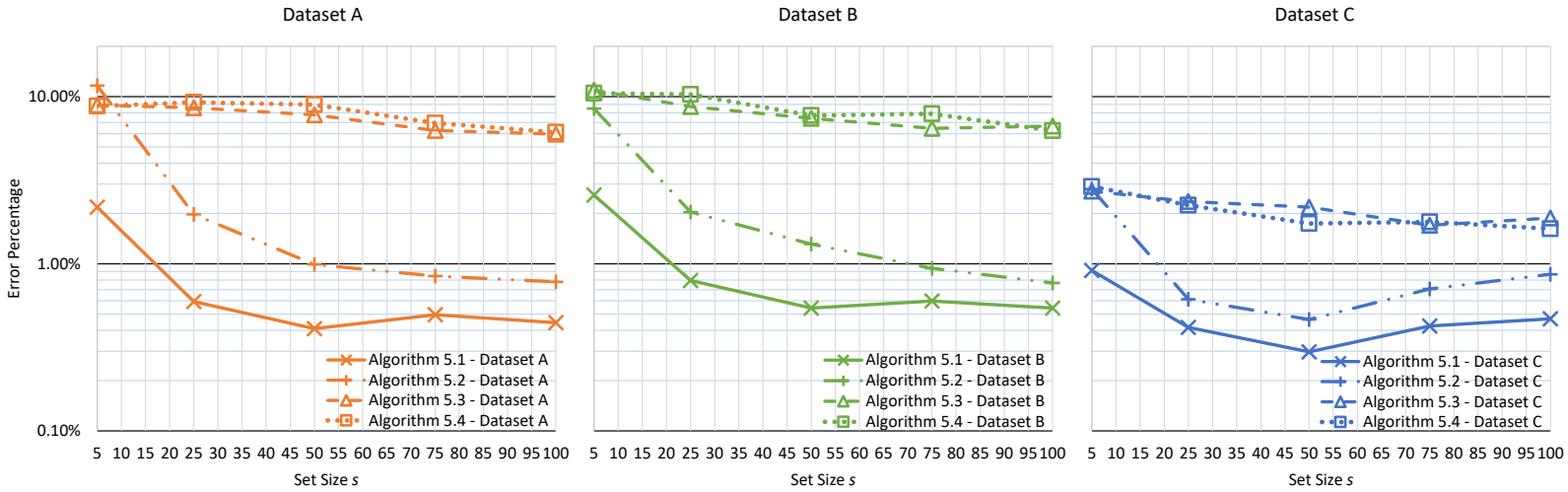

Figure 5.9: Algorithms Simulation Results – Mismatched Output

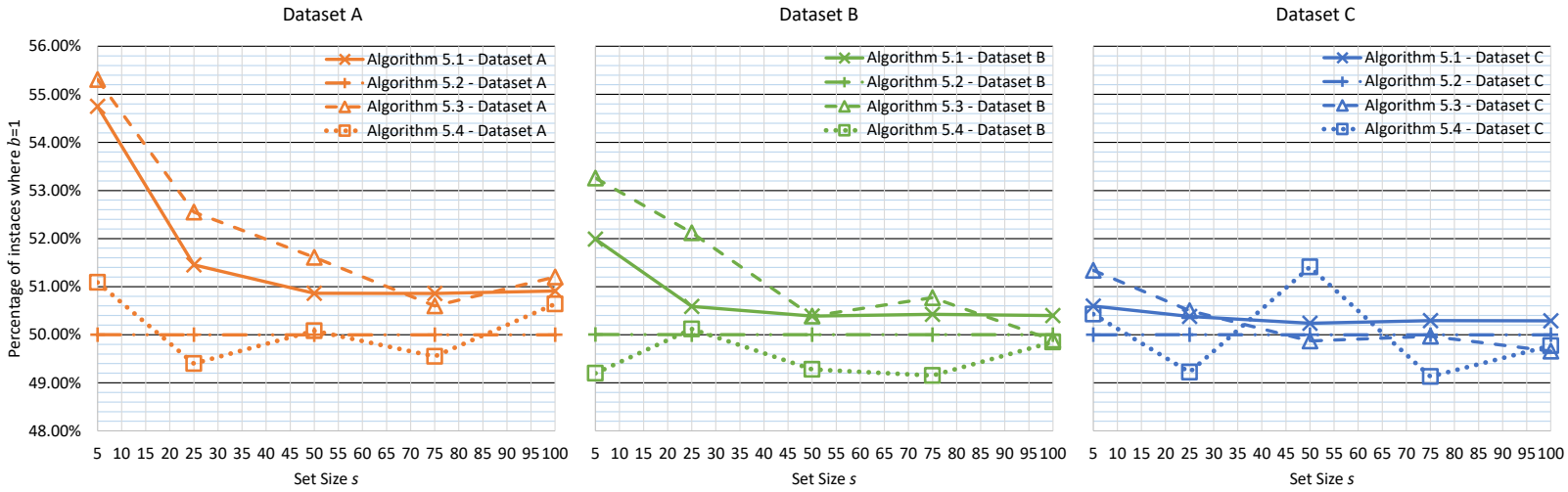

Figure 5.10: Algorithms Simulation Results – Percentage of $b = 1$ Output (Averaged Across II and V Signals)

5.6.2. Frequency Hopping Simulation Results

The results of frequency hopping simulation executed based on outputs of Algorithm 5.1 are shown in Figure 5.11. Note that the errors between the transmitter and the receiver were measured at three different points: 1) when b is in its original binary form; 2) after b is concatenated into a new 5-bit number; and 3) after the randomly generated data has gone through the frequency hopping system and received by the receiver.

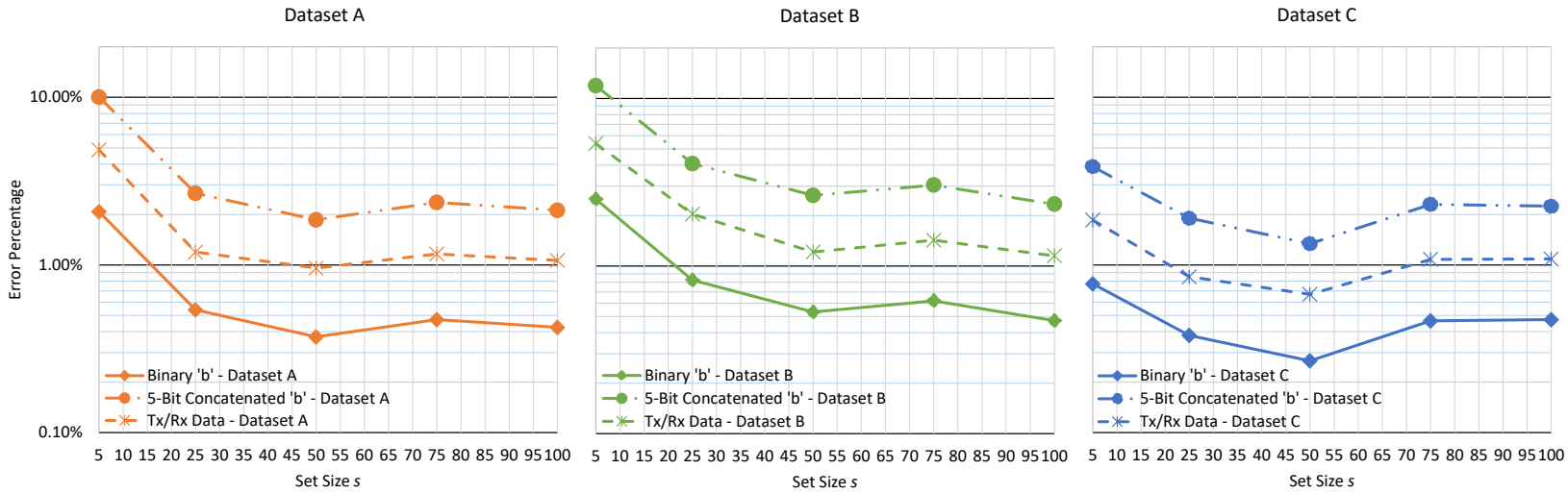

Figure 5.11: Frequency Hopping Simulation Results Based on Outputs of Algorithm 5.1

5.6.3. Performance Evaluation and Security Analysis

From the simulation results, Algorithm 5.1 achieves the best performance out of the four proposed algorithms and for all three datasets. Algorithm 5.2 performs quite poorly for datasets A and B at a relatively lower value of the comparison set size s, but improves significantly for a higher value of s. Algorithm 5.3 and Algorithm 5.4 perform worse than the other two in general and appear to be quite sensitive to dataset changes, with dataset C performing significantly better than the other two datasets for all values of s.

From the security perspective, let us consider two scenarios relating to an attacker's ability to predict the value of output b. The first is when the attacker can visually or otherwise observe physical activities carried out by the devices wearer/implantee. Such observations may assist in predicting the output of the algorithms, especially those that use consecutively accumulated IPIs like Algorithm 5.1 and Algorithm 5.2. The value of set size s used may also affect the likelihood of such predictions being successfully performed. The second scenario is when the attacker wildly guesses the potentially skewed output, as discussed in Section 5.6.1. There are ways to mitigate this and make sure the overall chance of $b = 0$ and $b = 1$ is approximately equal (i.e., 50%); for example, lines 6 and 8 of Algorithm 5.1 could be made the reverse of each other depending on an additional condition check on if $i > j$. Doing so would assist in reducing the probability of guessing the output in cases where the attacker does not know the pseudorandom determinant parameter k. However, if the attacker knows both the algorithm and the parameter k, they would still be able to gain an advantage by brute forcing the slightly more likely output that would be returned when $x \geq y$, especially for cases where a relatively lower value of comparison set size s is used.

In terms of other potential improvements, all the datasets used from the MIMIC-III database have a sampling rate of only 125 samples per second. If higher resolution

and more accurate data are available, there would be a potential that the output mismatch rate between two devices would decrease further for all algorithms. There would also likely be fewer instances of x equal to y in the comparisons which would mitigate the slightly skewed output towards $b = 1$ issue discussed in Section 5.6.1.

5.7. Conclusion

Four alternative algorithms are proposed, simulated and analysed in this chapter. The algorithms can be used with IPI biometrics to add another layer of protection to a traditional pseudorandomly determined frequency hopping pattern. In general, it was found that Algorithm 5.1 achieves the best performance out of the four and has the potential in being used to assist in mitigating jamming attacks. The value of comparison set size s should also be carefully selected to achieve the best balance of performance and output predictability. The availability of datasets with higher sampling rates would potentially improve the performance further, and might bring Algorithm 5.3 and Algorithm 5.4 back into consideration for their potential in offering better security in mitigating output predictability for cases where an attacker can visually or otherwise observe physical activities of the devices wearer/implantee.

Chapter 6 Conclusion and Future Works

Three studies have been comprehensively presented in Chapter 3, Chapter 4 and Chapter 5 of this thesis, where novel security solutions have been developed to improve security for the Internet of Things (IoT) in three different domains, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). This chapter summarises the works carried out, as well as the contributions made to the literature. Additionally, the remaining challenges identified within the studies, and the future research directions, are also discussed.

6.1. Research Summary

With the technological advancement in communications and computational systems, comes the consequences, one of which is the ever-increasing number of interconnected IoT devices which are becoming an integral part of human lives. Although this phenomenon brings about tremendous benefits to humanity, it also has associated challenges, one interesting area of which is security. The research works carried out in this thesis focus on improving security in three different IoT domains, namely, connected and autonomous vehicles (CAVs), Internet of Flying Things (IoFT), and human body interface and control systems (HBICS). These studies are comprehensively presented in Chapter 3, Chapter 4 and Chapter 5 of this thesis, and are summarised below.

In the area of CAVs, a novel scheme is proposed to facilitate secure and conditional privacy-preserving vehicular pseudonym issuance and management in a multi-jurisdictional road network. The architecture leverages the increasingly mature

permissioned consortium blockchain technology, the predicted wide availability of RSUs, and the highly viable, flexible, and well-established PKI technology. A small-scale simulation of the proposed architecture was successfully carried out using the Veins platform for integrated traffic and network simulation services, the Hyperledger Fabric platform for blockchain services, and the OpenSSL platform for PKI services. Simulation results indicate the feasibility of practical deployment of the scheme.

In the area of IoFT, a state-of-the-art intelligent Sybil attack detection scheme is proposed. The scheme uses supervised machine learning carried out on two physical layer features of the radio signals emitted from UAV nodes, namely, the received signal strength difference (RSSD) and the time difference of arrival (TDoA). Since the scheme functions using intrinsically generated physical layer data, it is less susceptible to various attacks commonly carried out on the upper layers, such as data spoofing. Moreover, no additional communications overheads of the UAV nodes are required for the functionality of this scheme. Simulations carried out using the OMNeT++/INET simulator and the Weka machine learning workbench platform revealed a high correct classification accuracy of above 91% on average, even for smart malicious nodes with power control capability operating at power levels not directly trained.

In the area of HBICS, a new frequency hopping approach is proposed, which uses IPI biometrics to add another layer of protection to the traditional pseudorandom frequency hopping system. MATLAB/Simulink simulations were carried out on the four different proposed alternative algorithms. Simulation results reveal the feasibility for some of the algorithms to be used. Potential improvements include the use of datasets with higher sampling rates.

Although the research works captured in this thesis fill knowledge gaps in the literature in several different areas, there are still open issues that should be addressed as future works, as listed in Section 6.3. As indicated in various places throughout this thesis, IoT is a broad area of study, consisting of numerous different

domains. Consequently, there are also many other IoT security-related open problems and challenges out there still waiting for new solutions to be invented to address them.

6.2. Contributions

The main contributions of this thesis can be summarised by chapter as follows:

- Chapter 3: A permissioned consortium blockchain-based conditional privacy-preserving vehicular pseudonym issuance and management system is proposed to be paired with the traditional PKI-based cryptography system. The proposed system is the first to focus on integrated secured access and management support for vehicular pseudonym issuance and management in a multi-jurisdictional road network. In terms of the consensus mechanism, the architecture is also flexible, allowing for efficient protocols to be used, such as the CFT protocol used in the simulation. The scheme also addresses the identified shortfalls of existing works from the perspective of achieving a better balance between connectivity and storage requirements.
    - Results: The simulations carried out demonstrate the feasibility of practical deployment of the scheme.
- Chapter 4: An intelligent Sybil attack detection scheme is proposed for FANETs-based IoFT environment. The scheme is the first of its kind, where detection is done using supervised machine learning carried out on two physical layer features of the radio signals, namely, the received signal strength difference (RSSD) and the time difference of arrival (TDoA). This scheme uses only two ground nodes to monitor radio signals emitted from the UAVs, which is less than what traditional position localisation methods would have required. Because of the use of only intrinsically generated physical layer data, there are

no additional costs in terms of communications overheads. Similarly, it is also less susceptible to various attacks commonly carried out on the upper layers, such as data spoofing. Furthermore, the scheme has been designed to be smart and flexible, in that it can function in situations where classic malicious nodes with fixed power are used, as well as in situations where smart malicious nodes with power control capability are used to manipulate signal strength.

  - Results: The simulation results reveal that the proposed scheme can achieve a high correct classification accuracy of above 91% on average, even for smart malicious nodes with power control capability operating at power levels not directly trained.

- Chapter 5: A frequency hopping approach in HBICS is proposed, which uses inter-pulse interval (IPI) biometrics to add another layer of protection to the traditional pseudorandom frequency hopping system. This is the first time IPI biometrics is proposed to be used to add another layer of protection to the traditional pseudorandom frequency hopping system, which is fundamentally different to the existing IPI biometrics solutions used for authentication and encryption applications. Four alternative algorithms are proposed to determine the frequency to operate on at any point in time, given that the HBICS devices are time synchronised and share a mutually known pseudorandom determinant parameter.

  - Results: Simulation results reveal that Algorithm 5.1 achieves the best performance out of the four and has the potential in being used to assist in mitigating jamming attacks. This is followed by Algorithm 5.2, and less preferably, Algorithm 5.3 and Algorithm 5.4 which resulted in significantly higher error percentages.

6.3. Remaining Challenges and Future Research Directions

6.3.1. Matters Identified in Chapter 3

Several matters identified in Chapter 3 as to be considered prior to system deployment are summarised as follows. Firstly, the variability in demand for pseudonym issuance, especially their peaks, needs to be further characterised. Secondly, the blockchain transactions processing throughput limitation should be studied further to discover how the system would behave if such a limit is reached. Thirdly, further assessments need to be made on the potential requirement for parallel execution of RSU's tasks by different processes. Finally, other potential risks associated with the selected permissioned consortium blockchain platform should be thoroughly identified and appropriately mitigated.

6.3.2. Matters Identified in Chapter 4

Matters identified in Chapter 4 as to be considered prior to system deployment are summarised as follows. Firstly, the study assumes that the free space path loss propagation model holds true and that signals from other systems in the surrounding area are coordinated in such a way that results in negligible interference effects on the functionality of the system, such as through the use of orthogonal frequency-division multiplexing. Therefore, further assessments would need to be carried out on the effects of interference and structural blockages applicable at the physical location the system is planned to be deployed. Additionally, prior to system deployment, further studies need to also be carried out on the actual expected mobility patterns, node density levels, flying space dimensions, signal emission rates and transmit power range.

The study also identified several improvement opportunities that can be done as future works. Firstly, the study identified that it should be investigated if the performance of the scheme can be improved even further if additional and/or different attributes, apart from RSSD and TDoA ratios, are used. It was also identified that unsupervised machine learning may need to be considered as an extension to the scheme, to cater for situations where datasets for machine learning training are not easily obtainable. Similarly, the possibility of extending the scheme to cater for other attack types in FANETs should also be considered. Finally, it was identified that adaptation of the proposed scheme for other application scenarios, such as in VANETs environment, should be investigated.

6.3.3. Matters Identified in Chapter 5

In Chapter 5, it was identified that the ECG datasets from the MIMIC-III database used in the study have a sampling rate of only 125 samples per second. The availability of datasets with higher sampling rates would potentially improve the performance of all four proposed algorithms. Additionally, this might also bring the poorly performed Algorithm 5.3 and Algorithm 5.4 back into consideration for their potential in offering better security in mitigating output predictability.

6.3.4. The Future of Internet of Things Security

As can be seen throughout this thesis, security has been and is still a challenging aspect of IoT. Although this thesis proposes solutions to address some significant security problems, numerous other challenges still remain. It is also important to remember that IoT is a broad field of study, covering so many different application areas. There is a possibility that the solutions proposed in this thesis may also be applicable to

similar issues being faced in other IoT domains; however, some adaptation efforts would likely be required.

Due to IoT being a very active research field, technological advancement is happening very rapidly. It is important that developers consider potential security issues as an integral part of any system design process. Nevertheless, there will likely still be technologies that get developed with security issues being neglected. Additionally, progressions associated with certain technologies may also cause existing security mechanisms to become ineffective. Thus, new solutions would need to be developed in such situations, one prominent example of which is the quantum field.

As outlined in Section 2.4, quantum technology will fundamentally change communications systems and networks security in the near future. It will provide an enabling platform for strong security and high-performance computation, which would also potentially assist ML/DL applications. On the contrary, it will also render many currently used cryptographical security systems ineffective. Unsurprisingly, this significant field of study opens up enormous research opportunities in the area of IoT security.

References

[1] Y. Harbi, Z. Aliouat, S. Harous and A. Bentaleb, “A Review of Security in Internet of Things,” *Wireless Personal Communications,* vol. 108, pp. 235-344, 2019.

[2] M. Noor, M. Binti and W. Hassan, “Current research on Internet of Things (IoT) security: A survey,” *Computer Networks,* vol. 148, pp. 283-294, 2019.

[3] J. Y. Khan and M. R. Yuce, Internet of Things (IoT) Systems and Applications, Pan Stanford Publishing, 2019.

[4] N. Miloslavskaya and A. Tolstoy, “Internet of Things: information security challenges and solutions,” *Cluster Computing,* vol. 22, no. 1, pp. 103-119, 2019.

[5] Cisco, “Cisco Annual Internet Report (2018–2023) White Paper,” 9 March 2020. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html. [Accessed 2 March 2023].

[6] A. Qayyum, M. Usama, J. Qadir and A. Al-Fuqaha, “Securing Connected & Autonomous Vehicles: Challenges Posed by Adversarial Machine Learning and the Way Forward,” *IEEE Communications surveys and tutorials,* vol. 22, no. 2, pp. 998-1026, 2020.

[7] Z. Mahmood, Connected Vehicles in the Internet of Things Concepts, Technologies and Frameworks for the IoV, Springer International Publishing, 2020.

[8] P. K. Singh, S. K. Nandi and S. Nandi, “A tutorial survey on vehicular communication state of the art, and future research directions,” *Vehicular Communications,* vol. 18, 2019.

[9] N. Xia, H.-H. Chen and C.-S. Yang, “Radio Resource Management in Machine-to-Machine Communications-A Survey,” *IEEE Communications Surveys & Tutorials,* vol. 20, no. 1, pp. 791-828, 2018.

[10] J. Wang, J. Liu and N. Kato, “Networking and Communications in Autonomous Driving: A Survey,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 2, pp. 1243-1274, 2019.

[11] Z. Machardy, A. Khan, K. Obana and S. Iwashina, “V2X Access Technologies: Regulation, Research, and Remaining Challenges,” *IEEE Communications Surveys & Tutorials,* vol. 20, no. 3, pp. 1858-1877, 2018.

[12] M. Obaidat, M. Khodjaeva, J. Holst and M. B. Zid, “Security and Privacy Challenges in Vehicular Ad Hoc Networks,” in *Connected Vehicles in the Internet of Things Concepts, Technologies and Frameworks for the IoV*, Springer International Publishing, 2020, pp. 223-251.

[13] M. Arif, G. Wang, M. Zakirul Alam Bhuiyan, T. Wang and J. Chen, “A survey on security attacks in VANETs: Communication, applications and challenges,” *Vehicular Communications,* vol. 19, 2019.

[14] S. Zaidi, M. Atiquzzaman and C. T. Calafate, “Internet of Flying Things (IoFT): A Survey,” *Computer Communications,* vol. 165, pp. 53-74, 2021.

[15] D. F. Pigatto, M. Rodrigues, J. V. de Carvalho Fontes, A. S. R. Pinto, J. Smith and K. R. L. J. C. Branco, “The Internet of Flying Things,” in *Internet of Things A to Z:*

Technologies and Applications, Hoboken, New Jersey, John Wiley & Sons, Inc., 2018, pp. 529-561.

[16] A. Chriki, H. Touati, H. Snoussi and F. Kamoun, "FANET: Communication, mobility models and security issues," *Computer Networks,* vol. 163, p. 106877, 2019.

[17] N. Gao, L. Liang, D. Cai, X. Li and S. Jin, "Coverage Control for UAV Swarm Communication Networks: A Distributed Learning Approach," *IEEE Internet of Things Journal,* 2022.

[18] Q. Zhang, M. Jiang, Z. Feng, W. Li, W. Zhang and M. Pan, "IoT Enabled UAV: Network Architecture and Routing Algorithm," *IEEE Internet of Things Journal,* vol. 6, no. 2, pp. 3727-3742, 2019.

[19] Y. Liu, J. Nie, X. Li, S. H. Ahmed, W. Y. B. Lim and C. Miao, "Federated Learning in the Sky: Aerial-Ground Air Quality Sensing Framework With UAV Swarms," *IEEE Internet of Things Journal,* vol. 8, no. 12, pp. 9827-9837, 2021.

[20] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith and A. Jamalipour, "Wireless Body Area Networks: A Survey," *IEEE Communications Surveys & Tutorials,* vol. 16, no. 3, pp. 1658-1686, 2014.

[21] S. Seneviratne, Y. Hu, T. Nguyen, G. Lan, S. Khalifa, K. Thilakarathna, M. Hassan and A. Seneviratne, "A Survey of Wearable Devices and Challenges," *IEEE Communications Surveys & Tutorials,* vol. 19, no. 4, pp. 2573-2620, 2017.

[22] K. Sowjanya and M. Dasgupta, "Survey of Symmetric and Asymmetric Key Management Schemes in the context of IoT based Healthcare System," in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, 2020.

[23] C. Camara, P. Peris-Lopez and J. E. Tapiador, “Security and privacy issues in implantable medical devices: A comprehensive survey,” *Journal of biomedical informatics,* vol. 55, pp. 272-289, 2015.

[24] H. Fotouhi, A. Causevic, K. Lundqvist and M. Björkman, “Communication and Security in Health Monitoring Systems -- A Review,” in *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, 2016.

[25] R. Cavallari, F. Martelli, R. Rosini, C. Buratti and R. Verdone, “A Survey on Wireless Body Area Networks: Technologies and Design Challenges,” *IEEE Communications Surveys & Tutorials,* vol. 16, no. 3, pp. 1635-1657, 2014.

[26] M. Usman, M. R. Asghar, I. S. Ansari and M. Qaraqe, “Security in Wireless Body Area Networks: From In-Body to Off-Body Communications,” *IEEE Access,* vol. 6, pp. 58064-58074, 2018.

[27] K. Karmakar, S. Saif, S. Biswas and S. Neogy, “WBAN Security: study and implementation of a biological key based framework,” in *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*, 2018.

[28] T. Yaqoob, H. Abbas and M. Atiquzzaman, “Security Vulnerabilities, Attacks, Countermeasures, and Regulations of Networked Medical Devices—A Review,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 4, pp. 3723-3768, 2019.

[29] O. Yousuf and R. Mir, “A survey on the Internet of Things security,” *Information and Computer Security,* vol. 27, no. 2, pp. 292-323, 2019.

[30] M. S. Ali, M. Vecchio, M. Pincheira, K. Dolui, F. Antonelli and M. H. Rehmani, “Applications of Blockchains in the Internet of Things: A Comprehensive

Survey," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 2, pp. 1676-1717, 2019.

[31] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 3, pp. 2671-2701, 2019.

[32] B. B. Gupta and A. Tewari, A beginner's guide to Internet of things security : attacks, applications, authentication, and fundamentals, Boca Raton : CRC Press, 2020.

[33] G. Ramezan, C. Leung and Z. J. Wang, "A Survey of Secure Routing Protocols in Multi-Hop Cellular Networks," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 4, pp. 3510-3541, 2018.

[34] R. W. van Der Heijden, S. Dietzel, T. Leinmuller and F. Kargl, "Survey on Misbehavior Detection in Cooperative Intelligent Transportation Systems," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 1, pp. 779-811, 2019.

[35] D. Eckhoff and I. Wagner, "Privacy in the Smart City-Applications, Technologies, Challenges, and Solutions," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 1, pp. 489-516, 2018.

[36] K. Zhao and L. Ge, "A Survey on the Internet of Things Security," *2013 Ninth International Conference on Computational Intelligence and Security,* pp. 663-667, 2013.

[37] Q. Jing, A. V. Vasilakos, J. Wan, J. Lu and D. Qiu, "Security of the Internet of Things: perspectives and challenges," *Wireless networks,* vol. 20, no. 8, pp. 2481-2501, 2014.

[38] I. Butun, P. Osterberg and H. Song, “Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures,” *IEEE Communications surveys and tutorials,* vol. 22, no. 1, pp. 616-644, 2020.

[39] P. Nespoli, D. Papamartzivanos, F. Gomez Marmol and G. Kambourakis, “Optimal Countermeasures Selection Against Cyber Attacks: A Comprehensive Survey on Reaction Frameworks,” *IEEE Communications Surveys & Tutorials,* vol. 20, no. 2, pp. 1361-1396, 2018.

[40] P. H. Rettore, G. Maia, L. A. Villas and A. A. F. Loureiro, “Vehicular Data Space: The Data Point of View,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 3, pp. 2392-2418, 2019.

[41] H. Khelifi, S. Luo, B. Nour, H. Moungla, Y. Faheem, R. Hussain and A. Ksentini, “Named Data Networking in Vehicular Ad Hoc Networks: State-of-the-Art and Challenges,” *IEEE Communications Surveys & Tutorials,* vol. 22, no. 1, pp. 320-351, 2020.

[42] R. Hussain and S. Zeadally, “Autonomous Cars: Research Results, Issues, and Future Challenges,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 2, pp. 1275-1313, 2019.

[43] Z. El-Rewini, K. Sadatsharan, D. F. Selvaraj, S. J. Plathottam and P. Ranganathan, “Cybersecurity challenges in vehicular communications,” *Vehicular Communications,* vol. 23, 2020.

[44] J.-H. Cho, D. P. Sharma, H. Alavizadeh, S. Yoon, N. Ben-Asher, T. J. Moore, D. S. Kim, H. Lim and F. F. Nelson, “Toward Proactive, Adaptive Defense: A Survey on Moving Target Defense,” *IEEE Communications Surveys & Tutorials,* vol. 22, no. 1, pp. 709-745, 2020.

[45] X. Wang, Z. Ning, M. Zhou, X. Hu, L. Wang, Y. Zhang, F. R. Yu and B. Hu, "Privacy-Preserving Content Dissemination for Vehicular Social Networks: Challenges and Solutions," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 2, pp. 1314-1345, 2019.

[46] N. Chen, M. Wang, N. Zhang and X. Shen, "Energy and Information Management of Electric Vehicular Network: A Survey," *IEEE Communications surveys and tutorials,* vol. 22, no. 2, pp. 967-997, 2020.

[47] S. Sharma and B. Kaushik, "A survey on internet of vehicles: Applications, security issues & solutions," *Vehicular Communications,* vol. 20, 2019.

[48] D. Manivannan, S. S. Moni and S. Zeadally, "Secure authentication and privacy-preserving techniques in Vehicular Ad-hoc NETworks (VANETs)," *Vehicular Communications,* vol. 25, 2020.

[49] A. Boualouache, S.-M. Senouci and S. Moussaoui, "A Survey on Pseudonym Changing Strategies for Vehicular Ad-Hoc Networks," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 1, pp. 770-790, 2018.

[50] I. Ali, A. Hassan and F. Li, "Authentication and privacy schemes for vehicular ad hoc networks (VANETs): A survey," *Vehicular Communications,* vol. 16, pp. 45-61, 2019.

[51] J. Sun, W. Wang, Q. Da, L. Kou, G. Zhao, L. Zhang and Q. Han, "An Intrusion Detection Based on Bayesian Game Theory for UAV Network," in *11th EAI International Conference on Mobile Multimedia Communications*, Qingdao, 2018.

[52] M. A. Lopez, M. Baddeley, W. T. Lunardi, A. Pandey and J.-P. Giacalone, "Towards Secure Wireless Mesh Networks for UAV Swarm Connectivity:

Current Threats, Research, and Opportunities," in *2021 17th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2021.

[53] K. Xing, S. S. R. Srinivasan, M. J. ". Rivera, J. Li and X. Cheng, "Attacks and countermeasures in sensor networks: A survey," in *Network Security*, Boston, Springer, 2010, pp. 251-272.

[54] J.-P. A. Yaacoub, M. Noura, H. N. Noura, O. Salman, E. Yaacoub, R. Couturier and A. Chehab, "Securing internet of medical things systems: Limitations, issues and recommendations," *Future Generation Computer Systems,* vol. 105, pp. 581-606, 2020.

[55] M. Rushanan, A. D. Rubin, D. F. Kune and C. M. Swanson, "SoK: Security and Privacy in Implantable Medical Devices and Body Area Networks," in *2014 IEEE Symposium on Security and Privacy*, 2014.

[56] M. R. K. Naik and P. Samundiswary, "Wireless body area network security issues — Survey," in *2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2016.

[57] V. Mainanwal, M. Gupta and S. K. Upadhayay, "A survey on wireless body area network: Security technology and its design methodology issue," in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2015.

[58] Manorama and I. Snigdh, "Anonymity in Body Area Sensor Networks-an Insight," in *2018 IEEE World Symposium on Communication Engineering (WSCE)*, 2018.

[59] B. Narwal and A. K. Mohapatra, “A survey on security and authentication in wireless body area networks,” *Journal of Systems Architecture,* vol. 113, p. 101883, 2021.

[60] L. Stabellini and M. M. Parhizkar, “Experimental Comparison of Frequency Hopping Techniques for 802.15.4-based Sensor Networks,” in *UBICOMM 2010 - 4th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2010.

[61] A. Sawand, S. Djahel, Z. Zhang and F. Naït-Abdesselam, “Multidisciplinary approaches to achieving efficient and trustworthy eHealth monitoring systems,” in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*, 2014.

[62] M. Kompara and M. Hölbl, “Survey on security in intra-body area network communication,” *Ad Hoc Networks,* vol. 70, pp. 23-43, 2018.

[63] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” in *Proceedings of the 6th ACM conference on Computer and communications security (CCS '99)*, 1999.

[64] D. K. Altop, A. Levi and V. Tuzcu, “Deriving cryptographic keys from physiological signals,” *Pervasive and Mobile Computing,* vol. 39, pp. 65-79, 2017.

[65] C. C. Poon, Y.-T. Zhang and S.-D. Bao, “A novel biometrics method to secure wireless body area sensor networks for telemedicine and m-health,” *IEEE Communications Magazine,* vol. 44, no. 4, pp. 73-81, 2006.

[66] A. Rizwan, A. Zoha, R. Zhang, W. Ahmad, K. Arshad, N. A. Ali, A. Alomainy, M. A. Imran and Q. H. Abbasi, “A Review on the Role of Nano-Communication in

Future Healthcare Systems: A Big Data Analytics Perspective," *IEEE Access,* vol. 6, pp. 41903-41920, 2018.

[67] F. Dressler and F. Kargl, "Towards security in nano-communication: Challenges and opportunities," *Nano communication networks,* vol. 3, no. 3, pp. 151-160, 2012.

[68] W. Zhang, T. Qin, M. Mekonen and W. Wang, "Wireless Body Area Network Identity Authentication Protocol Based on Physical Unclonable Function," in *2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, 2018.

[69] M. R. K. Naik and P. Samundiswary, "Survey on Game Theory Approach in Wireless Body Area Network," in *2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*, 2019.

[70] T. Salman, M. Zolanvari, A. Erbad, R. Jain and M. Samaka, "Security Services Using Blockchains: A State of the Art Survey," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 1, pp. 858-880, 2019.

[71] J. Xie, H. Tang, T. Huang, F. R. Yu, R. Xie, J. Liu and Y. Liu, "A Survey of Blockchain Technology Applied to Smart Cities: Research Issues and Challenges," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 3, pp. 2794-2830, 2019.

[72] G. Pulkkis, J. Karlsson and M. Westerlund, "Blockchain-Based Security Solutions for IoT Systems," in *Internet of things A to Z : technologies and applications*, Q. F. Hassan, Ed., Wiley, 2018, pp. 255-273.

[73] M. Saad, J. Spaulding, L. Njilla, C. Kamhoua, S. Shetty, D. Nyang and D. Mohaisen, "Exploring the Attack Surface of Blockchain: A Comprehensive

Survey," *IEEE Communications surveys and tutorials,* vol. 22, no. 3, pp. 1977-2008, 2020.

[74] T. Alladi, V. Chamola, N. Sahu and M. Guizani, "Applications of blockchain in unmanned aerial vehicles: A review," *Vehicular Communications,* vol. 23, 2020.

[75] B. Cao, Y. Li, L. Zhang, L. Zhang, S. Mumtaz, Z. Zhou and M. Peng, "When Internet of Things Meets Blockchain: Challenges in Distributed Consensus," *IEEE Network,* vol. 33, no. 6, pp. 133-139, 2019.

[76] F. Hussain, R. Hussain, S. A. Hassan and E. Hossain, "Machine Learning in IoT Security: Current Solutions and Future Challenges," *IEEE Communications surveys and tutorials,* vol. 22, no. 3, pp. 1686-1721, 2020.

[77] L. Zhang, D. Mu, W. Hu and Y. Tai, "Machine-Learning-Based Side-Channel Leakage Detection in Electronic System-Level Synthesis," *IEEE network,* vol. 34, no. 3, pp. 44-49, 2020.

[78] D. Karagiannis and A. Argyriou, "Jamming attack detection in a pair of RF communicating vehicles using unsupervised machine learning," *Vehicular Communications,* vol. 13, pp. 56-63, 2018.

[79] H. El-Sayed, H. A. Ignatious, P. Kulkarni and S. Bouktif, "Machine learning based trust management framework for vehicular networks," *Vehicular Communications,* vol. 25, 2020.

[80] Q. Mao, F. Hu and Q. Hao, "Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 4, pp. 2595-2621, 2018.

[81] Y. Yu, H. Li, R. Chen, Y. Zhao, H. Yang and X. Du, “Enabling Secure Intelligent Network with Cloud-Assisted Privacy-Preserving Machine Learning,” *IEEE Network,* vol. 33, no. 3, pp. 82-87, 2019.

[82] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali and M. Guizani, “A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security,” *IEEE Communications surveys and tutorials,* vol. 22, no. 3, pp. 1646-1685, 2020.

[83] H. Yao, P. Gao, P. Zhang, J. Wang, C. Jiang and L. Lu, “Hybrid Intrusion Detection System for Edge-Based IIoT Relying on Machine-Learning-Aided Detection,” *IEEE Network,* vol. 33, no. 5, pp. 75-81, 2019.

[84] M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, “Deep Learning for IoT Big Data and Streaming Analytics: A Survey,” *IEEE Communications Surveys & Tutorials,* vol. 20, no. 4, pp. 2923-2960, 2018.

[85] M. Usama, J. Qadir, A. Al-Fuqaha and M. Hamdi, “The Adversarial Machine Learning Conundrum: Can the Insecurity of ML Become the Achilles' Heel of Cognitive Networks?,” *IEEE Network,* vol. 34, no. 1, pp. 196-203, 2020.

[86] J. Qiu, L. Du, Y. Chen, Z. Tian, X. Du and M. Guizani, “Artificial Intelligence Security in 5G Networks: Adversarial Examples for Estimating a Travel Time Task,” *IEEE vehicular technology magazine,* vol. 15, no. 3, pp. 95-100, 2020.

[87] L. Gyongyosi, S. Imre and H. V. Nguyen, “A Survey on Quantum Channel Capacities,” *IEEE Communications Surveys & Tutorials,* vol. 20, no. 2, pp. 1149-1205, 2018.

[88] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis and P. Fan, “6G Wireless Networks: Vision, Requirements, Architecture, and Key

Technologies," *IEEE Vehicular Technology Magazine,* vol. 14, no. 3, pp. 28-41, 2019.

[89] J. Sliwa, "Environment, People, and Time as Factors in the Internet of Things Technical Revolution," in *Internet of things A to Z : technologies and applications*, Q. F. Hassan, Ed., Wiley, 2018, pp. 51-76.

[90] W. Saad, M. Bennis and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems," *IEEE network,* vol. 34, no. 3, pp. 134-142, 2020.

[91] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini and G. Bianchi, "Quantum Internet: Networking Challenges in Distributed Quantum Computing," *IEEE Network,* vol. 34, no. 1, pp. 137-143, 2020.

[92] J. Petit, F. Schaub, M. Feiri and F. Kargl, "Pseudonym Schemes in Vehicular Networks: A Survey," *IEEE Communications Surveys & Tutorials,* vol. 17, no. 1, pp. 228-255, 2015.

[93] Z. Lu, G. Qu and Z. Liu, "A Survey on Recent Advances in Vehicular Network Security, Trust, and Privacy," *IEEE transactions on intelligent transportation systems,* vol. 20, no. 2, pp. 760-776, 2019.

[94] L. Zhang, Q. Wu, J. Domingo-Ferrer, B. Qin and C. Hu, "Distributed Aggregate Privacy-Preserving Authentication in VANETs," *IEEE Transactions on Intelligent Transportation Systems,* vol. 18, no. 3, pp. 516-526, 2017.

[95] X. Lin, X. Sun, P.-H. Ho and X. Shen, "GSIS: A Secure and Privacy-Preserving Protocol for Vehicular Communications," *IEEE Transactions on Vehicular Technology,* vol. 56, no. 6, pp. 3442-3456, 2007.

[96] B. Rajendran, “Evolution of PKI ecosystem,” in *2017 International Conference on Public Key Infrastructure and its Applications (PKIA)*, Bangalore, India, 2017.

[97] K. Emara, W. Woerndl and J. Schlichter, “CAPS: context-aware privacy scheme for VANET safety applications,” in *Proceedings of the 8th ACM Conference on security & privacy in wireless and mobile networks*, New York, 2015.

[98] S. Bao, . Y. Cao, A. Lei, P. Asuquo, H. Cruickshank, Z. Sun and M. Huth, “Pseudonym Management Through Blockchain: Cost-Efficient Privacy Preservation on Intelligent Transportation Systems,” *IEEE Access,* vol. 7, pp. 80390-80403, 2019.

[99] I. Makhdoom, M. Abolhasan, J. Lipman, R. P. Liu and W. Ni, “Anatomy of Threats to the Internet of Things,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 2, pp. 1636-1675, 2019.

[100] V. Ortega, F. Bouchmal and J. F. Monserrat, “Trusted 5G Vehicular Networks: Blockchains and Content-Centric Networking,” *IEEE Vehicular Technology Magazine,* vol. 13, no. 2, pp. 121-127, 2018.

[101] N. B. Truong, K. Sun, G. M. Lee and Y. Guo, “GDPR-Compliant Personal Data Management: A Blockchain-Based Solution,” *IEEE Transactions on Information Forensics and Security,* vol. 15, pp. 1746-1761, 2020.

[102] R. Gupta, S. Tanwar, F. Al-Turjman, P. Italiya, A. Nauman and S. W. Kim, “Smart Contract Privacy Protection Using AI in Cyber-Physical Systems: Tools, Techniques and Challenges,” *IEEE Access,* vol. 8, pp. 24746-24772, 2020.

[103] H. Honar Pajooh, M. Rashid, F. Alam and S. Demidenko, “Hyperledger Fabric Blockchain for Securing the Edge Internet of Things,” *Sensors (Basel, Switzerland),* vol. 21, no. 2, p. 359, 2021.

[104] M. Belotti, N. Božić, G. Pujolle and S. Secci, “A Vademecum on Blockchain Technologies: When, Which, and How,” *IEEE Communications Surveys & Tutorials,* vol. 21, no. 4, pp. 3796-3838, 2019.

[105] Hyperledger, “Private data,” 2020. [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.3/private-data/private-data.html. [Accessed 9 March 2023].

[106] L. Benarous, B. Kadri and A. Bouridane, “Blockchain-Based Privacy-Aware Pseudonym Management Framework for Vehicular Networks,” *Arabian journal for science and engineering,* vol. 45, no. 8, pp. 6033-6049, 2020.

[107] A. Lei, H. Cruickshank, Y. Cao, P. Asuquo, C. P. Anyigor Ogah and Z. Sun, “Blockchain-Based Dynamic Key Management for Heterogeneous Intelligent Transportation Systems,” *IEEE Internet of Things Journal,* vol. 4, no. 6, pp. 1832-1843, 2017.

[108] N. Malik, P. Nanda, A. Arora, X. He and D. Puthal, “Blockchain Based Secured Identity Authentication and Expeditious Revocation Framework for Vehicular Networks,” in *18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, New York, 2019.

[109] N. Lasla, M. Younis, W. Znaidi and D. B. Arbia, “Efficient Distributed Admission and Revocation Using Blockchain for Cooperative ITS,” in *9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, Paris, France, 2018.

[110] S. Guo, X. Hu, Z. Zhou, X. Wang, F. Qi and L. Gao, “Trust access authentication in vehicular network based on blockchain,” *China Communications,* vol. 16, no. 6, pp. 18-30, 2019.

[111] J. Ma, T. Li, J. Cui, Z. Ying and J. Cheng, “Attribute-Based Secure Announcement Sharing Among Vehicles Using Blockchain,” *IEEE Internet of Things Journal,* vol. 8, no. 13, pp. 10873-10883, 2021.

[112] C. Sommer, “Veins,” 2021. [Online]. Available: https://veins.car2x.org/. [Accessed 9 March 2023].

[113] The Linux Foundation, “Hyperledger Fabric – Hyperledger,” 2022. [Online]. Available: https://www.hyperledger.org/use/fabric. [Accessed 9 March 2023].

[114] Eclipse Foundation, “Eclipse SUMO - Simulation of Urban MObility,” 2023. [Online]. Available: https://www.eclipse.org/sumo/. [Accessed 9 March 2023].

[115] OpenSim Ltd., “OMNeT++ Discrete Event Simulator,” 2022. [Online]. Available: https://omnetpp.org/. [Accessed 9 March 2023].

[116] OpenSSL Software Foundation, “OpenSSL Cryptography and SSL/TLS Toolkit,” OpenSSL Software Foundation, 2023. [Online]. Available: https://www.openssl.org/. [Accessed 9 March 2023].

[117] Hyperledger, “What’s new in Hyperledger Fabric v2.x,” 2020. [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.3/whatsnew.html#private-data-enhancements. [Accessed 9 March 2023].

[118] D. Luzeaux, J.-R. Ruault and J.-L. Wippler, Complex Systems and Systems of Systems Engineering, London: ISTE, 2011.

[119] G. S. Parnell, P. J. Driscoll and D. L. Henderson, Decision Making in Systems Engineering and Management, Hoboken, New Jersey: Wiley, 2011.

[120] W. Li, M. Cao, Y. Wang, C. Tang and F. Lin, “Mining Pool Game Model and Nash Equilibrium Analysis for PoW-Based Blockchain Networks,” *IEEE Access,* vol. 8, pp. 101049-101060, 2020.

[121] Q. Gu, T. Fan, F. Pan and C. Zhang, “A vehicle-UAV operation scheme for instant delivery,” *Computers & Industrial Engineering,* vol. 149, p. 106809, 2020.

[122] J. Zheng, R. Chen, T. Yang, X. Liu, H. Liu, T. Su and L. Wan, “An Efficient Strategy for Accurate Detection and Localization of UAV Swarms,” *IEEE Internet of Things Journal,* vol. 8, no. 20, pp. 15372-15381, 2021.

[123] S. Murali and A. Jamalipour, “A Lightweight Intrusion Detection for Sybil Attack Under Mobile RPL in the Internet of Things,” *IEEE Internet of Things Journal,* vol. 7, no. 1, pp. 379-388, 2020.

[124] A. Arshad, Z. Mohd Hanapi, S. Subramaniam and R. Latip, “A survey of Sybil attack countermeasures in IoT-based wireless sensor networks,” *PeerJ Computer Science,* vol. 7, pp. e673-e673, 2021.

[125] E. Frank, M. A. Hall and I. H. Witten, “The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques",” in *Data Mining: Practical Machine Learning Tools and Techniques*, 4 ed., Morgan Kaufmann, 2016.

[126] D. Dardari, E. Falletti and M. Luise, Satellite and Terrestrial Radio Positioning Techniques: A signal processing perspective, Academic Press, 2012.

[127] D. Munoz, F. Bouchereau, C. Vargas and R. Enriquez, Position Location Techniques and Applications, Academic Press, 2009.

[128] X. Li, Z. D. Deng, L. T. Rauchenstein and T. J. Carlson, "Contributed Review: Source-localization algorithms and applications using time of arrival and time difference of arrival measurements," *Review of Scientific Instruments,* vol. 87, no. 4, p. 041502, 2016.

[129] A. Vasudeva and M. Sood, "Survey on sybil attack defense mechanisms in wireless ad hoc networks," *Journal of network and computer applications,* vol. 120, pp. 78-118, 2018.

[130] A. K. Singh, "Innovative survey of defense machinery against sybil attacks over wireless ad-hoc network on IoT," *Journal of Engg. Research,* vol. 9, no. 2, pp. 92-105`, 2021.

[131] G. Shobana and X. A. R. Arockia, "Detection Mechanism on Vehicular Adhoc Networks (VANETs) A Comprehensive Survey," *International Journal of Computer Science & Network Security,* vol. 21, no. 6, pp. 294-303, 2021.

[132] Y. Zhang, B. Das and F. Qiao, "Sybil Attack Detection and Prevention in VANETs: A Survey," in *Proceedings of the Future Technologies Conference (FTC) 2020*, Vancouver, 2020.

[133] N. C. Velayudhan and A. Anitha, "Sybil Attack in VANET Operating in an Urban Environment: An Overview," in *Advances in Communication Systems and Networks*, Singapore, Springer, 2020, pp. 433-442.

[134] B. Hammi, Y. M. Idir, S. Zeadally, R. Khatoun and J. Nebhen, "Is It Really Easy to Detect Sybil Attacks in C-ITS Environments: A Position Paper," *IEEE Transactions on Intelligent Transportation Systems,* 2022.

[135] M. Kabbur and V. A. Kumar, “MAR_Sybil: Cooperative RSU Based Detection and Prevention of Sybil Attacks in Routing Process of VANET,” *Journal of physics. Conference series,* vol. 1427, no. 1, p. 12009, 2020.

[136] Y. Yuan, L. Huo, Z. Wang and D. Hogrefe, “Secure APIT Localization Scheme Against Sybil Attacks in Distributed Wireless Sensor Networks,” *IEEE Access,* vol. 6, pp. 27629-27636, 2018.

[137] S. Lv, X. Wang, X. Zhao and X. Zhou, “Detecting the Sybil Attack Cooperatively in Wireless Sensor Networks,” in *2008 International Conference on Computational Intelligence and Security*, 2008.

[138] S. Abbas, M. Merabti, D. Llewellyn-Jones and K. Kifayat, “Lightweight Sybil Attack Detection in MANETs,” *IEEE Systems Journal,* vol. 7, no. 2, pp. 236-248, 2013.

[139] A. Angappan, T. P. Saravanabava, P. Sakthivel and K. S. Vishvaksenan, “Novel Sybil attack detection using RSSI and neighbour information to ensure secure communication in WSN,” *Journal of Ambient Intelligence and Humanized Computing,* vol. 12, p. 6567–6578, 2021.

[140] C. F. E. de Melo, T. Dapper e Silva, F. Boeira, J. M. Stocchero, A. Vinel, M. Asplund and E. P. de Freitas, “UAVouch: A Secure Identity and Location Validation Scheme for UAV-Networks,” *IEEE Access,* vol. 9, pp. 82930-82946, 2021.

[141] E. Walia, V. Bhatia and G. Kaur, “Detection Of Malicious Nodes in Flying Ad-HOC Networks (FANET),” *International Journal of Electronics and Communication Engineering,* vol. 5, no. 9, pp. 6-12, 2018.

[142] A. Jamalipour and S. Murali, “A Taxonomy of Machine-Learning-Based Intrusion Detection Systems for the Internet of Things: A Survey,” *IEEE Internet of Things Journal,* vol. 9, no. 12, pp. 9444-9466, 2022.

[143] I. H. Sarker, A. I. Khan, Y. B. Abushark and F. Alsolami, “Internet of Things (IoT) Security Intelligence: A Comprehensive Overview, Machine Learning Solutions and Research Directions,” *Mobile Networks and Applications,* 2022.

[144] U. Farooq, N. Tariq, M. Asim, T. Baker and A. Al-Shamma'a, “Machine learning and the Internet of Things security: Solutions and open challenges,” *Journal of Parallel and Distributed Computing,* vol. 162, pp. 89-104, 2022.

[145] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen and L. Hanzo, “Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks,” *IEEE Communications Surveys & Tutorials,* vol. 22, no. 3, pp. 1472-1514, 2020.

[146] H. T. Friis, “A Note on a Simple Transmission Formula,” *Proceedings of the IRE,* vol. 34, no. 5, pp. 254-256, 1946.

[147] OpenSim Ltd., “INET Framework,” 2022. [Online]. Available: https://inet.omnetpp.org/. [Accessed 9 March 2023].

[148] OpenSim Ltd., “Showcases > Mobility > 3D Mobility,” [Online]. Available: https://inet.omnetpp.org/docs/showcases/mobility/spatial/doc/. [Accessed 9 March 2023].

[149] OpenSim Ltd., “Tutorials > Wireless Tutorial > Step 13. Configuring a more accurate path loss model,” [Online]. Available: https://inet.omnetpp.org/docs/tutorials/wireless/doc/step13.html. [Accessed 9 March 2023].

[150] OpenSim Ltd., “Tutorials > Wireless Tutorial > Step 14. Introducing antenna gain,” [Online]. Available: https://inet.omnetpp.org/docs/tutorials/wireless/doc/step14.html. [Accessed 9 March 2023].

[151] E. Anthi, L. Williams, M. Słowińska, G. Theodorakopoulos and P. Burnap, “A Supervised Intrusion Detection System for Smart Home IoT Devices,” *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 9042-9053, 2019.

[152] V. Sharma, A. P. S. Chouhan and D. Bisen, “Prediction of activation energy of biomass wastes by using multilayer perceptron neural network with Weka,” *Materials Today: Proceedings,* vol. 57, pp. 1944-1949, 2022.

[153] J. R. Quinlan, C4.5: Programs for Machine Learning, San Mateo, California: Morgan Kaufmann, 1993.

[154] I. H. Witten, E. Frank, M. A. Hall and C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, 4 ed., Morgan Kaufmann, 2017.

[155] Y. Wang and I. H. Witten, “Induction of model trees for predicting continuous classes,” in *Poster papers of the 9th European Conference on Machine Learning*, Prague, 1997.

[156] R. C. Holte, “Very Simple Classification Rules Perform Well on Most Commonly Used Datasets,” *Machine Learning,* vol. 11, p. 63–90, 1993.

[157] W. W. Cohen, “Fast Effective Rule Induction,” in *Proceedings of the Twelfth International Conference on Machine Learning*, Tahoe City, California, 1995.

[158] J. Fürnkranz and G. Widmer, “Incremental Reduced Error Pruning,” in *Proceedings of the Eleventh International Conference, Rutgers University*, New Brunswick, NJ, 1994.

[159] U. Uludag, S. Pankanti, S. Prabhakar and A. K. Jain, “Biometric cryptosystems: issues and challenges,” *Proceedings of the IEEE,* vol. 92, no. 6, pp. 948-960, 2004.

[160] B. Gopalakrishnan and M. A. Bhagyaveni, “Anti-jamming communication for body area network using chaotic frequency hopping,” *Healthcare Technology Letters,* vol. 4, no. 6, pp. 233-237, 2017.

[161] A. E. W. Johnson , T. J. Pollard, L. Shen, L.-W. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi and R. G. Mark, “MIMIC-III, a freely accessible critical care database,” *Scientific Data,* vol. 3, no. 1, pp. 160035-160035, 2016.

[162] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,” *Circulation,* vol. 101, no. 23, pp. E215-E220, 2000.

[163] X. Fan and Z. Tan, “Simulink Implementation of Frequency-hopping Communication System and Follower Jamming,” in *2018 IEEE International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*, 2018.