QATAR UNIVERSITY

COLLEGE OF ENGINEERING

MULTI-LAYER ATTACK DETECTION MODEL FOR BLUETOOTH-CONNECTED DEVICES IN SMART HEALTHCARE SYSTEM

BY

MOHAMMED ZUBAIR

A Thesis Submitted to

the College of Engineering

in Partial Fulfillment of the Requirements for the Degree of

Masters in Computing

January 2023

COMMITTEE PAGE

The members of the Committee approve the Thesis of Mohammed Zubair defended on 15/01/2023.

Dr. Abdulla Khalid A M Al-Ali
Thesis Supervisor

Dr. Devrim Unal
Thesis Co-Supervisor

Dr. Loay Sabry Ismail
Committee Member

Dr. Mahmoud Barhamgi
Committee Member

Dr. Roberto Di Pietro
Committee Member

Approved:

Dr. Khalid Kamal Naji, Dean, College of Engineering

ABSTRACT

Zubair, Mohammed, Masters: January: 2023, Masters in Computing

Title: Multi-Layer Attack Detection Model for Bluetooth-connected Devices in Smart Healthcare System.

Supervisor of Thesis: Abdulla Khalid A M Al-Ali, Devrim Unal.

Internet of Things (IoT) is an interconnected network of heterogeneous things through the Internet. The current and next generation of smart healthcare systems are dependent on Internet of Medical things (IoMT) devices (e.g, smart wireless medical sensors). In the current interconnected world, Bluetooth technology plays a vital role in short-range of communication due to its less resource consumption due to its flexibility and low resource consumption which suits the IoMT architecture and design. Smart health system present an ever-expanding attack surface due to the continuous adoption of a broad variety of Internet of Medical Things (IoMT) devices and applications. IoMT is a common approach to smart city solutions that deliver long-term benefits to critical infrastructure such as smart healthcare. As smart healthcare applications rely on distributed control optimization, artificial intelligence (AI), and in particular, deep learning (DL), offers an effective approach to mitigate cyber-attacks.

In this thesis we presents a decentralized, predictive DL-based process to autonomously detect and block malicious traffic and provide end-to-end defense against network attacks in IoMT devices. Furthermore, we provide the *BlueTack* dataset for Bluetooth-based attacks against IoMT networks. To the best of our knowledge this is the first intrusion detection dataset for the Bluetooth Classic and Bluetooth Low Energy (BLE). Using the BlueTack dataset, we devise a multi-layer intrusion detection method that uses deep-learning techniques. Then, we propose a decentralized architecture for deploying this IDS on the edge device of a smart-healthcare system that may be deployed in a smart city. The presented multi-layer intrusion detection models achieve performance in the range of 97%-99.5% based on the F1 scores.

DEDICATION

To my beloved wife.

ACKNOWLEDGMENTS

Alhamdulilah! All praise to Allah who grant me endurance and strength to complete this study.

I would like to express my gratitude to my Thesis supervisors, Dr. Abdulla Al-Ali and Dr. Devrim Unal for encouraging me during my master studies.

This Thesis has received funding from Hamad Medical Corporation Ambulance Service and was made possible by an NPRP grant, NPRP 10-0125-170250 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Internet of Things (IoT) interconnects and integrate physical objects through internet in diverse field such as smart healthcare, smart home, industries. The utilization of IoT devices are expected to reach 75 billion by the end of 2025 [1]. IoT technology is transforming our cities into smart cities. Smart cities use technologies for sensing, networking, and computation to enhance the quality of life and well-being of their inhabitants. Such smart cities also require a new service-centric computing paradigm for the next generation network (5G, 6G, and beyond) [2].

While there are numerous networking technologies available for long-range communication, the most widely used technology for close proximity communications is Bluetooth. Bluetooth is well suited to operations on resource-constrained mobile devices due to its low power consumption, low cost, and support for multimedia such as data and audio streaming. Bluetooth is also widely used in smart healthcare systems to enable untethered wireless communication between smart healthcare devices. Recently, Bluetooth was prominent in its adoption for contact-tracing applications in the fight against the COVID-19 global pandemic [3]. By the year 2030 [4], the number of IoT devices is expected to surge by 124 billion. Besides, the healthcare economy statistics predict that the market for IoT devices will grow from 20 billion dollars in 2015 to 70 billion dollars in 2025. It is also reported that 30.3% of the IoT devices in use are in the health sector **Healthcare (2020)**. The massive usage of IoT devices in heterogeneous networks provides various services using multiple technologies and protocols (such as Wi-Fi, Long Term Evolution (LTE), Bluetooth, ZigBee) make the task of securing such networks very complex. Information Systems Audit and Control Association (ISACA) research [5] on smart cities identified the security of IoT devices as an important focus area as numerous smart city critical infrastructure (CI) concepts (e.g., intelligent transport, healthcare system, and energy distribution) rely on the robustness and security of smart technologies and IoT devices [6].

As the number of Internet of Medical Things (IoMT) devices increases, the network becomes congested, which leads to bandwidth and latency bottlenecks [7]. For instance, an IoMT device sends data to a medical professional for analysis regularly. This transmission of data to the cloud can potentially cause latency and bandwidth congestion in the communication path, which could endanger the life of the patient. To address this challenge, the edge cloud concept has emerged for the IoMT paradigm. An edge cloud improves efficiency and provides more reliability for the smart healthcare system. The quick response time and reduced energy consumption result in longer battery life for medical devices and reduce the usage of network bandwidth.

The exponential growth of IoT devices and the massive interconnectivity between such devices however greatly opens up the potential attack surface for smart healthcare services that may be exploited by malicious actors. IoT devices are vulnerable to various medium- and high-severity attacks [8]. Various vulnerabilities allow the intruders to perform a wide range of attacks such as Denial of Service (DoS), Distributed DoS (DDoS), Man-in-the-Middle (MITM), data leakage, and spoofing. These attacks result in the unavailability of system resources and can lead to physical harm to the individuals when the patient is ambulance-bound or hospital-bound. According to a report of the Global Connected Industries cybersecurity, 82% of healthcare facilities experience cyberattacks, amongst which 30% target IoT devices [8]. The potential weakness in the network, IoT device, and protocol allows the attackers to access the network completely in an unauthorized way (e.g., Mirai attack) [9]. Apart from these cyber attacks, insecure operating systems, and application vulnerabilities are some of the other major threats to the healthcare system. Investigations show that 83% of the IoT devices run on outdated operating systems, and around 51% of the cyber threats in the health sector concern imaging devices, which lead to disruption of communication between patients and medical professionals. Moreover, 98% of the IoT device traffic is in plain text that can be intercepted by adversaries.

Traditional security mechanisms can not be enforced in the IoT network because the network protocol stack itself may have numerous vulnerabilities. Zero-day attacks are very difficult to be detected by traditional security mechanisms due to computational expenses, which do not go well with the resource-constrained nature of typical IoT devices [10]. Conventional perimeter security controls only defend against external attacks, but they fail to detect internal attacks within the network. An intelligent and faster detection mechanism is required to guarantee the security of the IoT network for countering the new threats before the network is compromised.

In this Thesis, our focus is on the security of Bluetooth communication in smart healthcare systems. After reviewing the significant security problems, we focus on the detection of wireless attacks against IoMT. Wireless attacks are performed when the data is at rest or in transmission from one device to another device in a wireless medium over different channels using various protocols namely Bluetooth Low Energy (BLE), Bluetooth Basic Rate/ Enhanced Data Rate (BR/EDR), Wi-Fi, Long Range (LoRA), etc. The openness of the wireless network poses threats to the entire network and can end up in the compromise of the entire system. The attacker may perform various attacks such as peer-to-peer, denial-of-service, eavesdropping, man-in-the-middle (MITM), and authentication attacks to take over the IoMT device or complete network.

Motivation

The security of IoMT objects is distinct from the sensors network and cyber-physcial-system (CPS). The exponential growth of IoT devices and the massive interconnectivity between such devices however greatly opens up the potential attack surface for smart healthcare services that may be exploited by malicious actors. The significant attention are paid by the security and researcher society all over the world towards intelligent smart devices. The potential weakness in the network, IoT device, and protocol allows the attackers to access the network completely in an unauthorized way .Apart from these cyber attacks, insecure operating systems and application vulnerabilities are some of the other major threats to the healthcare system. Investigations show that 83% of the IoT devices run on outdated operating systems, and around 51% of the cyber threats in the health sector concern imaging devices, which lead to disruption of communication between patients and medical professionals. Moreover, 98% of the IoT device traffic is in plain text that can be intercepted by adversaries. Traditional security mechanisms can not be enforced in the IoT network because the network protocol stack itself may have numerous vulnerabilities. An intelligent and faster detection mechanism is required to guarantee the security of the IoT network for countering the new threats before the network is compromised

Thesis Contributions

The main contributions of this study are as follows:

1. We have curated a novel first-of-its-kind *BlueTack dataset* for Bluetooth-based IoT attacks. The BlueTack Dataset consists of popular attacks against Bluetooth

BR/EDR protocol namely: Bluesmack, DoS, DDoS, and similar attacks such as DDoS and MITM attack on the BLE protocol. To the best of our knowledge, this is the first intrusion detection dataset for the Bluetooth classic protocol and BLE.

2. A secure and scalable framework for the deployment of an Intrusion Detection System (IDS) on the edge nodes of IoT-based healthcare systems in smart cities. The framework guarantees quicker identification of malicious activities to ensure the safety of critically ill patients transported by ambulance.

3. A multi-layer intrusion detection model using Deep Learning (DL) to protect the edge nodes of the smart healthcare IoMT system. Since IoMT is composed of several resource-constrained devices, deploying the DL model on the IoMT device itself for advanced functionality is impractical. Hence, The IDS is divided into two layers: *Layer_1* (where preprocessing is performed on IoMT devices or the edge node) and *Layer_2* (standalone GPU capability device on which the DL model is deployed). The proposed DL-based IDS achieves 99% accuracy while being deployed in a real-time scenario.

Thesis Overview

The structure of the rest of the Thesis is as follows: Chapter 2 provides background and related works, in which we discuss literature of Bluetooth technology, security of IoMT, Bluetooth protocol and IDS for Bluetooth-enables system. In Chapter 3, describes about a scalable architecture of smart healthcare system. In chapter 4, we explain our multi-layer attack detection model using deep learning. At last, the thesis is concluded with future research direction in chapter 5 respectively.

CHAPTER 2: BACKGROUND AND RELATED WORK

Bluetooth Technology

Bluetooth technology was invented in 1994 by L.M. Ericsson [11]. Initially this technology was referred as MCL (Multi-Communicator Links) and was dismissed due to its lower baud rate of 721 kbps. Later, after further developments it was approved by IEEE for 802.15.1. in 2002. Progressively, now it holds the market with respect to its specifications and data transfer data rate that has exceeded upto 20 Mbps. Basically Bluetooth has 3 different classes for connecting ranges: class 1 connectivity range is 100 meters and permitted power is about 100 mW, Class 2 connectivity range is 10 meters and permitted power is about 25 mW and Class 3 connectivity range is 1 meter and permitted power is about 1 mW. The strength of this technology is that it supports both data and audio (i.e, asynchronous and synchronous links) where Re-transmission of packets can be done through asynchronous link for error handling. The network is Ad-hoc in nature and known as piconets, where two or more Bluetooth devices are physically nearest to communicate on the same channel with same frequency hopping sequence. It operates on the unlicensed ISM band at 2.4GHz using advance spectrum frequency hopping technique, the hopping rate is about 1600 hop/sec of full duplex signal. Bluetooth chip induces a wavelength those are limited to some operating frequencies in a specified range (short range of communication). However, problems arises if same frequencies are used by many devices by causing signal interruptions or collisions [12]. In order to avoid and manage this issue, the signals are expanded over wide range of frequencies. So far, various protocols have been adopted in the Bluetooth standards (i.e, TCP/IP stack running over PPP), Bluetooth network encapsulation protocol (BNEP), object exchange protocol for exchanges (vCalendar and vCard) with IrDA interfaces. Similar to the internet protocol stack, Bluetooth also have a list of protocols that promote its communication. We will provide a brief overview on the Bluetooth protocol stack in the below subsection.

Bluetooth Protocol Stack

The protocol stack of Bluetooth shown in Figure 1 [13], is classified in to core specification and profile specification. Core specification, demonstrates the protocols from physical layer to the data link control with its management functions **9**. The profile specification deals with the different protocol and functions, and it delineates how to use Bluetooth technology **10**.

RFCOMM is a serial port emulator/cable replacement protocol and emulates the serial interface RS-232 standard. It is placed on the top of the L2CAP. This allows replacement of serial cable and enables the operation of distinct applications and protocols. The signalling, establishing and controlling of voice calls and data calls between Bluetooth devices is done by bit oriented protocol TCS-BIN (Telephony control protocol specification-binary) [14]. The host controllers interface (HCI) links the baseband and L2CAP, to access the hardware, control the register, render the command interface, and to link manager and baseband controller [15]. It can be identified as boundary of the hardware and software.

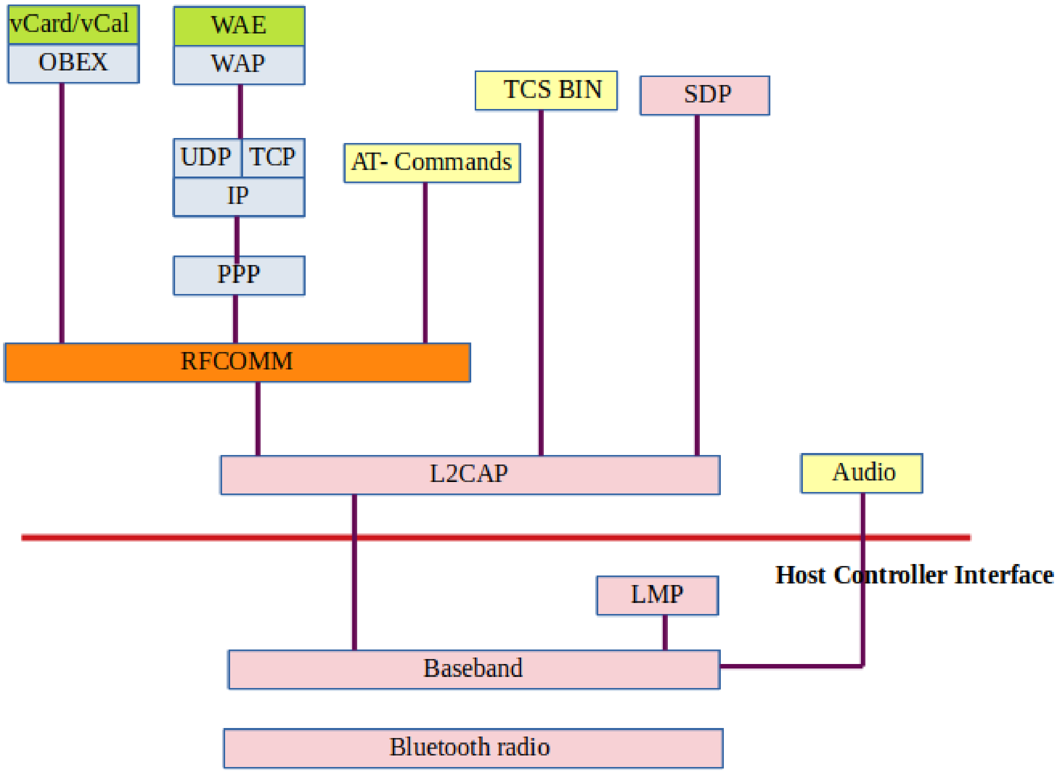

Figure 2.1: Protocol Stack of Bluetooth

Now, the protocol stack of Bluetooth is illustrated elaborately for further understanding.

1. **Bluetooth Radio Layer:** Carrier frequencies and output power are defined in this layer. For transmission, frequency hopping, time-division duplex scheme and Gaussian Frequency Shift Keying (GFSK) form modulation is used at the hopping rate 1600 hops/sec [16]. Each slot is defined as the time difference between two hops.

2. **Baseband layer:** This layer defines not only physical links and packet format but also perform frequency hopping and interference mitigation [17]. Time Division Duplex (TDD) is used for the transmission directions. Bluetooth packets at baseband layer consist of following three sections:

- Access code: The packet partition is used for time synchronization and piconet identification (i.e, channel, devices and inquiry access codes (CAC), (DAC), (IAC)). Access code constitutes of 4 bit preamble, 64-bit synchronization filed and 4 bit trailer.
- Packet header: It consists of packet type, packet flow control, error control, address and checksum. For each bit, sequence number (SEQN) and acknowledge number (ARQN) is used by alternating protocols [18]. Packet header protection is done by a header checksum with Forward error correction(FEC). Since, it holds valuable link and survive bit error.
- Payload: The structure of the payload field is depend on the type of link that is being used and usually, its size is upto 343 bytes.

3. **Physical Link:** Physical links are of two types of: Synchronous Connection-oriented Link (SCO) and Asynchronous Connectionless Link(ACL).

   (a) Synchronous connection-oriented link (SCO):

   In this link master device fixes two consecutive slots at fixed interval of times. Three SCO links are supported by master device for the same or different slave device. Two links are supported by a slave from different masters and three links from the same master device. Voice connection need symmetrical circuit switched, point-to-point connection. Various Forward Error Correction (FEC) schemes can be applied for increasing the data amount depending upon the channel error rate. Re-transmission of voice over data cannot be done in SCO link, A robust technique, continuous variable slope delta (CVSD) is applied for voice encoding to ensure the security of the data **haartsen1998bluetooth**.

(b) Asynchronous connectionless link (ACL):In this link, polling scheme is adopted by the master device and slave devices are addressed in the priority slots. Only one link is generated between a master and slave. Data transmission is carried out by various slots packets. Usually, in noisy environment, packet protection is accomplished by FEC schemes with high link error rate and overhead too. Hence, Bluetooth proposed a fast repeat request (ARQ) scheme for error free and efficient data transmission **haartsen1998bluetooth**. Various application requires symmetrical or asymmetrical, packet switched, point-to- multi point transfer scenarios which are supported by Bluetooth communication.

4. **Link Manager Protocol (LMP):** This protocol supervises different features of the radio connection between a master device and slave device along with their parameters specification and setting. It enhances the baseband functionality and covers Authentication, Encryption and many other functions. During piconet establishment the device begins inquiring by broadcasting an inquiry access code(IAC) to 32 wake-up carriers. Standby devices sniff the IAC messages periodically on the wake-up carrier and enter the inquiry mode. Once the device is detected the master starts to build a connection, from that point the device behaves as a salve. The device enters page mode, if the inquiry is successful and two different stages were defined later after this stage. By setting up a piconet, the master is able to communicate after identifying the device. The special hopping sequence is calculated by master based on address received by the devices and in return

synchronizes with the master clock, finally the devices enters the fully-connected states [19].

The connection states are categorized as Active state and the Low power states. In the Active state, the slave listens, transmits and receives by participating in the piconet. ACL and SCO links can be used for this purpose. All devices in the active state must have a 3-bit active member address (AMA). Bluetooth devices gets into the low power state for less power or battery consumption. Low power state is further classified into three, namely: Stiff, Hold and Park states.

- **Sniff state:** In this the devices are appointed to a limited number of slots for transmission to its slaves. But, the device retains its Active Member Address and consumes high power for transmission when compared to Hold and Park states.
- **Hold state:** At this state, the device block ACL transmission but does not release its AMA. Exchanging of SCO packets still goes on, if there is no activity in one piconet then the slave minimizes the power consumption or participates in another piconet.
- **Park State:** The device lowers its duty cycle in this state and least power is utilized for this process. The device discharges AMA and it receives a parked member address (PMA).

5. **L2CAP:** The logical link control and adaption protocol(L2CAP), it provides logical channel between multiple Bluetooth-enabled devices with QoS. In this layer **Connectionless, signalling and Connection-oriented** channels are available with own channel identifier (CID). The dedicated CID values 1,2 are used signalling

and connectionless channels respectively. The CID value greater than 64 is used by connection oriented channel for recognizing the connection uniquely. The protocol Data Units (PDUs) uses the CID values from 3 to 63, that deliver the operation of segmentation and reassembling.

6. **Service Discovery Protocol(SDP):** This protocol provides the new services (discovers) in the proximity of the Bluetooth signal range. It facilitates services to the SDP client by installing the SDP server on the device. The service records are maintained and identified by 32-bit service that posses unique value and ID.

Security of IoMT

IoMT devices perform diverse tasks in smart healthcare systems such as recording electrical impulses through Electrocardiogram (ECG) or monitoring blood glucose or blood pressure. For ambulance-bound patients, IoMT devices monitor the patient's activity, save critical information about the patient's physiological signals, and trigger alerts to the medical staff inside the ambulance or to a remote monitoring device through the cloud. As the complete information of the patient flows through the edge node (IoMT gateway), securing the IoMT attack surface assumes critical importance. An attacker may target the gateway to modify the patient vital information before sending to the medical practitioner or by performing denial of service (DoS) attack to disrupt the communication or service unavailable. These activities leads to risk the patient life. Rasool et al. [20] have reviewed various security issues of IoMT devices. The authors describe the vulnerabilities that exist in these devices which can be exploited by attackers easily. In our article, we are considering internal and external threats that are targeted against IoMT infrastructure. Since these devices are severely resource-constrained, it is easy to render these devices unavailable by draining their battery with devastating implications. Thus, our focus in this paper is on attacks that may drain the battery of these devices or which make the device unavailable due to multiple ping requests.

Communication in Smart Healthcare System

The typical architecture of a Smart healthcare system is shown in Fig. 2.2. A typical smart healthcare system comprises three domains: IoT domain, cloud domain, and user domain, which serve the role of generating data, storing data, and making diagnoses, respectively. The *IoT domain* consists of wireless medical devices, actuators, sensors, gateways, and other devices. Here, the focus is on acquiring patients' data from IoMT devices and transmitting it to the cloud for storage and subsequent access. The *cloud domain* is stratified by the edge and core cloud. The edge cloud is placed on the premises of the medical facility to ensure continuous connectivity and low latency, in addition to quicker diagnosis of acute cases. Core cloud provides massive storage and comprehensive analysis of data, and it helps in the diagnosis of current symptoms based on previous related records. The *user domain* delivers the processed data from other domains to the authorized clinical staff. Integration and streaming of vast volumes of data from different sources are visualized in various forms such as graphics, images, tabular, and other representations.

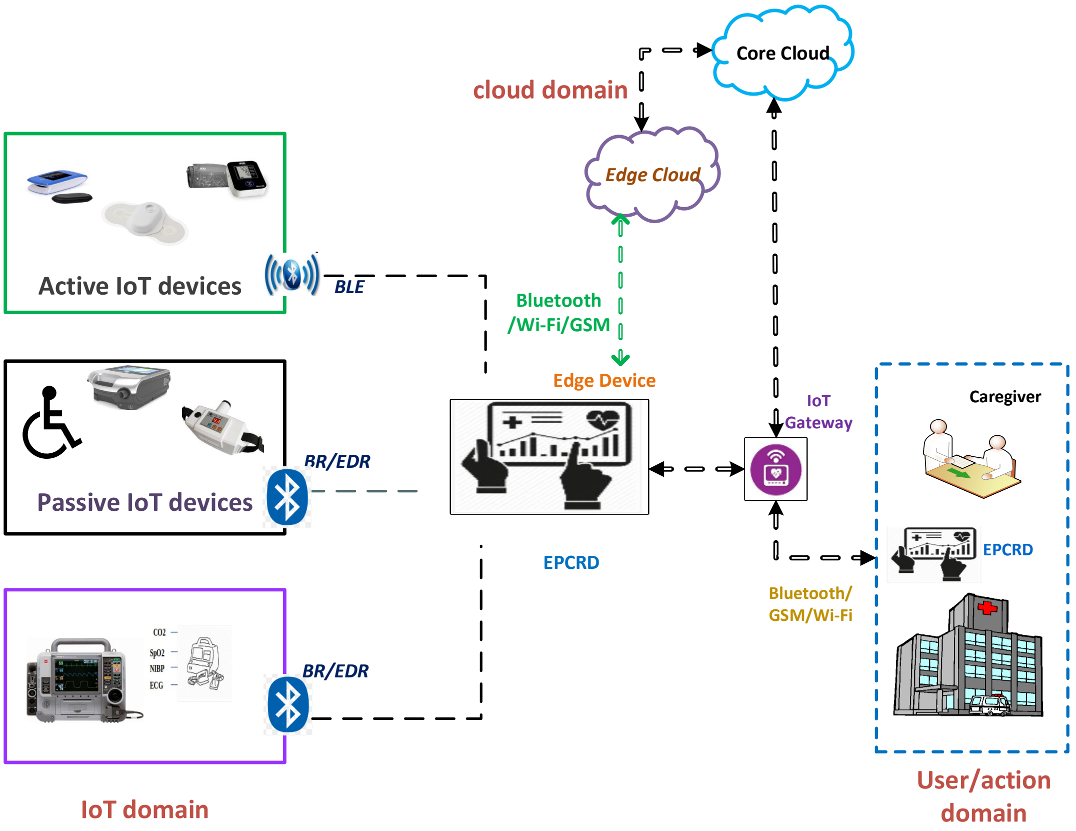

Figure 2.2: The use of Bluetooth and related protocols (BLE: Bluetooth Low Energy; BR/EDR: Bluetooth Basic Rate/Enhanced Data Rate) in a typical smart healthcare system for communication between Electronic Patient Care Record Device (EPCRD) and other entities over the Edge and the Cloud.

Medical devices (such as defibrillators and insulin pumps) that are continuously linked with the patient for medical treatment are referred to *Active Medical Devices* (AMD). On the other hand, medical devices (such as home monitoring devices and medical beds) whose focus is on periodic monitoring of the patient physical condition and report generation are called *Passive Medical Devices* (PMD). Wireless communication technologies are adopted for communication in IoT devices such as Near Field Communication (NFC), RFID, Wi-Fi, Bluetooth, LTE, and LoRA. Various IoMT devices use different wireless technologies. Most of the AMD and PMD utilize Bluetooth Classic, and V4.X and V5. The Bluetooth technology provides a generic profile for the medical IoT devices to use

Table 2.1: Technical details of Bluetooth Technology

Features	**Bluetooth Classic**	**Bluetooth V4.X (BLE)**	**Bluetooth V5 (BLE)**
Medium-access technique (MAC)	Frequency hopping	Frequency hopping	Frequency hopping
Multihop solution	Yes	Yes	(Yes)
Network topology	Piconet, scatternet	Star-bus mesh	Star-bus mesh
Radio frequency	2.4GHz	2.4GHz	(2.4GHz)
Nominal data rate (Mb/s)	1-3	1	2
Distance-range (meters)	Up to 100	Up to 100	Up to 200
Latency(ms)	less than 100	less than 6	less than 3
Nodes/slaves	7	Unlimited	Unlimited
Message-Size (bytes)	Up to 358	31	255

the 2.4 GHz frequency band, as recommended by the International Telecommunication Unit (ITU) [21]. Some of the basic differences between BR/EDR and the BLE are showcased in Table 2.1.

It is noted that Bluetooth-enabled devices have two modes of operation. In the single mode, a BLE device can not interface with a device that is operating on BR/EDR, and vice versa. Whereas in dual-mode, both BR/EDR and BLE devices can communicate with each other. However, the major concern is about security and privacy in all Bluetooth versions. In this paper, we focus on the detection of attacks against the BR/EDR and BLE, since the medical sensor and data collection devices in the considered testbed utilize this version of Bluetooth.

An Electronic Patient Care Record Device (EPCRD) collects health records and transmits them to the cloud. Edge computing allows and enables the technologies for computation at the edge of the healthcare network. It accomplishes the tasks of caching, processing storage, computation offloading, request distribution, and delivery of the services from the cloud end to the user end. In our proposed approach, we leverage edge cloud technology and deploy the IDS on the edge device/nodes of the smart healthcare system.

Vulnerabilities in the Bluetooth Protocols

The major factor of vulnerability in Bluetooth devices is the version that is being used for communication. Table [22] in the Appendix describes the vulnerabilities and security flaws of Bluetooth devices for different versions [23]. Few of the known vulnerabilities are identified by the researchers, such as MITM, Bluesmack, Battery Drain attacks, and Backdoor attack [24]. Recently, researchers identified the "SweynTooth" vulnerability affecting implantable medical devices (e.g., insulin pumps, pacemakers, blood glucose monitors) and hospital equipment (e.g., patient monitors and ultrasound machines) that work on BLE [25]. Bluetooth protocol has problems due to the encryption key length and improper storage of the link keys can be potentially manipulated by the adversary [9].

Intrusion Detection System

Some prior research on Intrusion detection systems (IDS) dedicated to the cyber-physical system [26] or smart environments using the Wi-Fi protocol against DoS attack [27] has adopted various AI techniques, such as ML and DL. One such approach [28] framework, a hybrid model that is based on the Principal Component Analysis (PCA) and Information Gain (IG) incorporating the Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), and instance-based learning models to identify the intrusions in the network. The model is trained and tested using the NSL-KDD, Kyoto 2006+, and ISCX 2012 dataset, and the optimal features are selected using an ensemble classifier. However, the performance of the model is evaluated with some publicly available datasets, which are not real-time datasets.

Sawarna et al. [29] have proposed an efficient IDS based on Deep Neural Network (DNN) using Principle Component Analysis - Grey Wolf Optimization (PCA-GWO), it eliminates the adversarial activities by providing faster alerts. This research is conducted to address the problem of data dimensionality for publicly available huge datasets. They tested the NSL-KDD dataset on various ML and DNN models to detect anomalies, among which the best accuracy is attained by the DNN. Baburaj et al. [30] have proposed a cloud based healthcare system using an SVM model to predict the health condition of a patient. The confidential data is accessed by a legitimate user only. This approach focused on data mining techniques using ML models, but not to identify the anomalies in the system.

Likewise, a supervised approach for detecting intrusions in IoT devices in a smart home is proposed by Eanthi et al. [31]. In this approach, a lightweight standalone three-layer IDS framework is built using a Decision Tree (DT) classifier with promising results. Nevertheless, the evaluation of the proposed model is based on a simulation performed on the open source Weka tool and the effectiveness of the IDS is not tested against real-time traffic and attacks.

IDS for Bluetooth Enabled Systems

Very few researchers have focused on the security perspective of Bluetooth technology, especially intrusion detection. Various attacks against Bluetooth devices are discussed below to emphasize the need for effective intrusion detection for Bluetooth-enabled medical IoT devices. Bluetooth technology provides a generic profile for the IoMT devices and it uses the 2.4 GHz frequency. It is identified as an attractive protocol for the healthcare system due to its robustness, lesser power consumption, low cost, suitability for shorter distance communication, and support for data and audio streaming. Moreover, it helps in the IoT domain for machine-to-machine (M2M) communication [32]. Compromising the IoMT devices could lead to sensitive patient information being revealed through the interception and decoding of the data and audio/video streaming packets. An IDS detects malicious activities or policy violations that bypass the security mechanism on a network and is the process of monitoring and detecting unauthorized events intruding on the network. An intruder is the one who escalates the privileges of the users to gain access to data or services or to control the entire network. Bluetooth-enabled systems require a different approach and standard IDS developed for other protocols are not effective due to the difference in traffic patterns and the highly constrained nature of Bluetooth devices [33].

Haataja et al. [34] have proposed a Bluetooth intrusion detection and prevention system based on a set of rules by investigating Bluetooth security to discover malicious communication on the Bluetooth network. Krzyszto et al. [35] have proposed a detection system to identify the malicious behavior of Bluetooth traffic in a Bluetooth mesh network. Multiple watchdog nodes are used for cooperative decisions in different areas of the mesh network. Malicious activities are detected based on the Received Signal Strength Indicator (RSSI). However, this model encountered the problem of modeling transmission range and RSSI parameters with obstacles such as furniture and walls.

This detection mechanism is not deployed to a variety of attacks and is evaluated in a simulated environment.

Similarly, Satam et al. [36] built a Bluetooth IDS (BIDS), where normal behavior of the Bluetooth traffic is defined based on the n-gram approach, and malicious traffic is classified using traditional ML algorithms. This method attained the highest precision of about 99.6% and recall of 99.6% against DoS attack. Yet, the effectiveness of the IDS is not tested against different datasets and other attacks. An anomaly-based intrusion detection system is proposed by Psatam et al. [37] to detect multiple attacks on the Bluetooth protocol using ML models by following the zero-trust principle. Nevertheless, the model is not tested using different attacks and datasets. Newaz et al. [38] have focused on the detection of the BLE for multiple attacks using ML models to identify the abnormal behavior of the BLE traffic from the normal traffic pattern. The evaluation of the model is done on their own real-time traffic for an ideal dataset, but has not been tested on other datasets.

From the above literature and Table 2.2, it is observed that the existing IDS approaches that are dedicated to healthcare IoT systems are at the initial stage of development. Few of the proposed IDS have validated their models on the data of the network simulation (dataset) or on a small number of IoT devices but they are not tested on multiple datasets. Moreover, these proposed IDS models detect malicious activities on the network by identifying the traffic pattern as normal or abnormal. It is also important to identify the various types of attacks on the network. In the below subsection, we describe the healthcare system in use by this paper and Bluetooth technology (BR/EDR and BLE) deployed.

Table 2.2: Various BIDS approaches in comparison to our proposed models. Our Bluetooth intrusion detection covers both Bluetooth classic and Bluetooth low-energy protocols.

Article	**Description**	**Model**	**Protocol**	**Data Availability**	**Problem Address**	**Deployed**
21	DNN based IDS using (PCA-GWO)	Deep neural network (DNN)	Ethernet / Wi-Fi	(NSL-KDD) -Publicly available	Data dimensionality and anomalies detection	No
22	Cloud based health-care system	Support vector machine (SVM)	-	(Vital) - No	Data mining techniques	No
24	Bluetooth IDS for Bluetooth network	Defined set of rules	Bluetooth	(BR/EDR) -No	Malicious traffic detection	Yes
25	Bluetooth mesh IDS-based on RSSI	Mesh-network	Bluetooth	(BR/EDR RSSI signals) -No	simulation and detection of malicious patterns	Yes
26	BIDS for IoT	ML-models	BR/EDR	(BR/EDR) -No	Malicious traffic based on n-gram	No
27	BIDS for IoT	ML-models	BR/EDR	(BR/EDR) -No	Multiple attack detections based on zero-trust	No
29	BLE IDS for medical devices	ML-Models	BLE	(BLE) - No	Multiple attack detections for irregular traffic flow	Yes
Our approach	Bluetooth IDS for health-care system	DL,ML-models	BR/EDR and BLE	(BR/EDR, BLE) - yes	Multiple attack detection of BR/EDR, BLE traffic	Yes

CHAPTER 3: SCALABLE ARCHITECTURE

By acknowledging the weakness of the security mechanism, we proposed and designed a scalable architecture to transfer vital information of a patient to the medical professional efficiently without alteration, tampering the patient data. The objective is to secure the Bluetooth communication against abnormal activities on the edge device of the smart healthcare system. The proposed architecture has enforced security policies, and detection mechanisms at the edge cloud and edge nodes to ensure fast response and secure emergency services. Edge computing helps to process the data efficiently with a quicker response time and assists with the deployment of the IDS. Figure 3.1 represents the proposed architecture of smart healthcare for detecting malicious behaviors of ambulance-bound, Bluetooth-enabled IoT medical devices in the smart healthcare system.

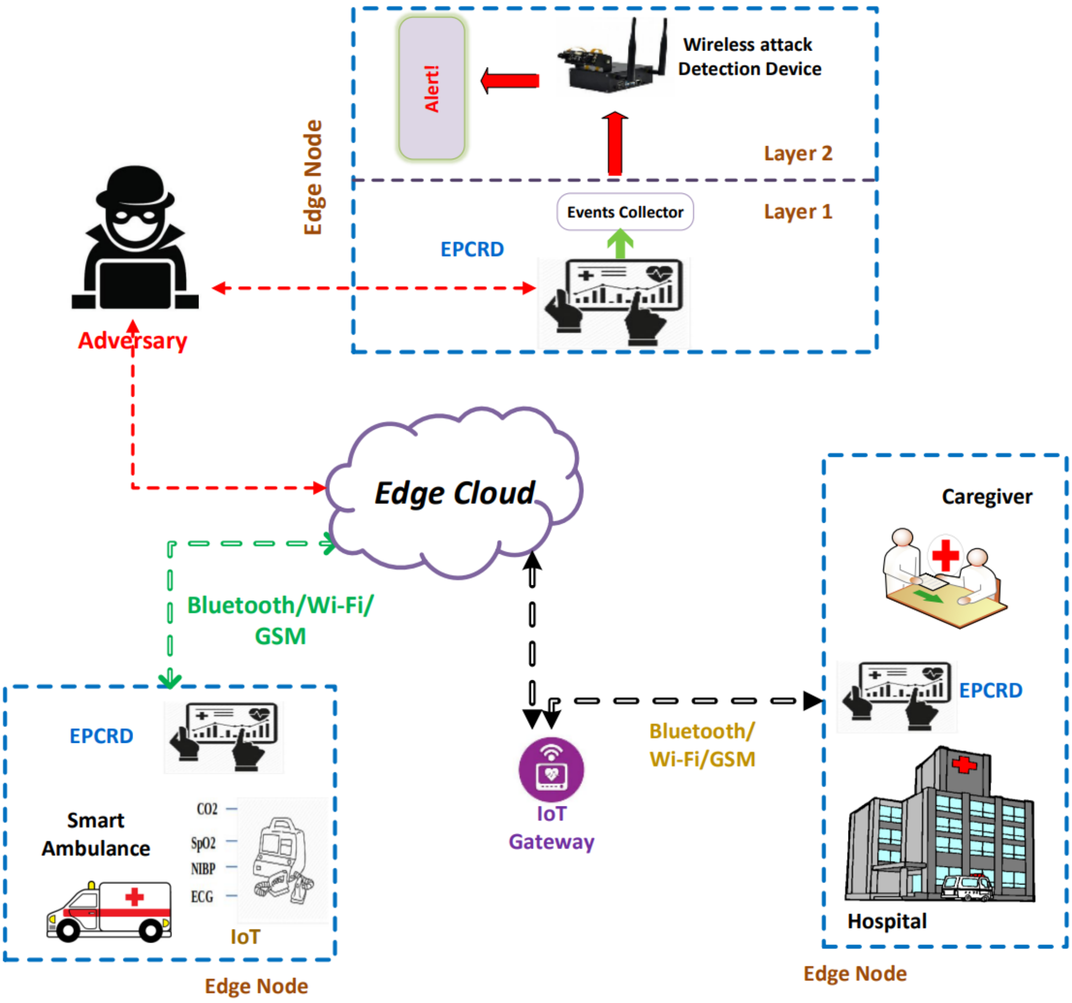

Figure 3.1: Architecture of the proposed security framework. The proposed system involves Edge cloud for reducing request and response delay. The IDS is multi-level to suit the resource restrictions of IoMT devices.

As the complete information of the patient flows in and out through the medical IoT gateway, it allows for a potential attacking surface to compromise the complete system by (1) targeting the medical IoMT gateway tampering the vital information of a patient before receives to the medical experts. or by (2) launching DoS/DDoS or MITM attacks to make the services unavailable or alter the information that leads the patient life at risk. To detect such abnormal activities while data transferring or data at rest , we enforce a multi-layer IDS on the edge device of the smart healthcare system. This IDS comprises two layers, namely, Layer_1 and Layer_2. Layer_1 accumulates patient vital information through IoMT-gateway and carry out data preprocess, feature engineering, selection of features by applying different ML algorithm. Layer_2 will detect the malicious activities of the Bluetooth traffic on the edge device/node using a DNN classifier. Next, we describe in detail the features of each layer:

*Layer*_1

On this layer, the data is collected for preprocessing, featuring engineering and for feature selection. Data preprocessing helps to provide the privacy of the medical information from the IoT devices because the information received from IoT devices is in plain text that can be intercepted by adversaries to perform medium- and high-severity attacks [39]. Data pre-processing is performed to transform actual data into data compatible with ML/DL models. For this process, we have used numericalization (where a string is converted into integer (stoi) then encoded into tokenized sentences before feeding to any model) and normalization. Data pre-processing helps the model to be trained and tested quickly. It also increases the accuracy of classification. We provide a detailed explanation of these stages below.

Eliminating / Dropping features: While capturing the traffic, we eliminated some information such as source and destination information. This choice is due to two major issues, firstly, in some scenarios, it is difficult for the sniffer to collect this information [40], while in other cases the adversary may spoof its address giving wrong information. In both cases, the classifier attempts to misclassify the traffic by replacing the missing values with some random numbers, giving higher False positives and True negatives. Likewise, we have eliminated some other unimportant and irrelevant features.

Feature selection: In this process, significant features are selected from the dataset by applying various feature selection techniques [41] [42]. Feature selection increases the model performance, decreases computational cost, and also increases storage efficiency. Additionally, using appropriate features reduce the problem of overfitting.

There are various ML approaches for selecting features, such as filter-based methods, wrapper methods, embedded, and statistical methods. In the univariate selection technique, a statistical test is applied to each feature to select the features which have a strong bond with the output variables. We used *Chi square (chi-2)*, in Eqn. 3.1 that gives the level of independence between the features x_t and the label y_t, it differentiate the chi-distribution with degree of freedom as 1.

$$\chi^2(x_t, y_t) = \frac{M.(FZ - PQ)^2}{(F+P)(F+Q)(P+Z)(Q+Z)} \tag{3.1}$$

where F indicates the frequency of the features and their labels in a dataset; P = frequency of the features emerges without a label; Q = frequency of label emerges without features; Z = frequency of neither features nor label emerges in the given dataset; and M = no. of training samples x_t = $x_1, x_2, ...x_i$ and prediction sequence y_t = $y_1, y_2, ...y_i$.

Recursive Feature Elimination (RFE), is an effective method to find an optimal set of features for both regression and classification tasks. Initially, it creates a model dependent on all the features and estimates the importance of each feature of a given dataset. It priorities the features based on the rank order and eliminates those features that are of least importance based on the evaluation metrics (in our case, we selected accuracy as a metric to find the optimal features) of the proposed model (DNN), which is depicted in the Fig. 3.2

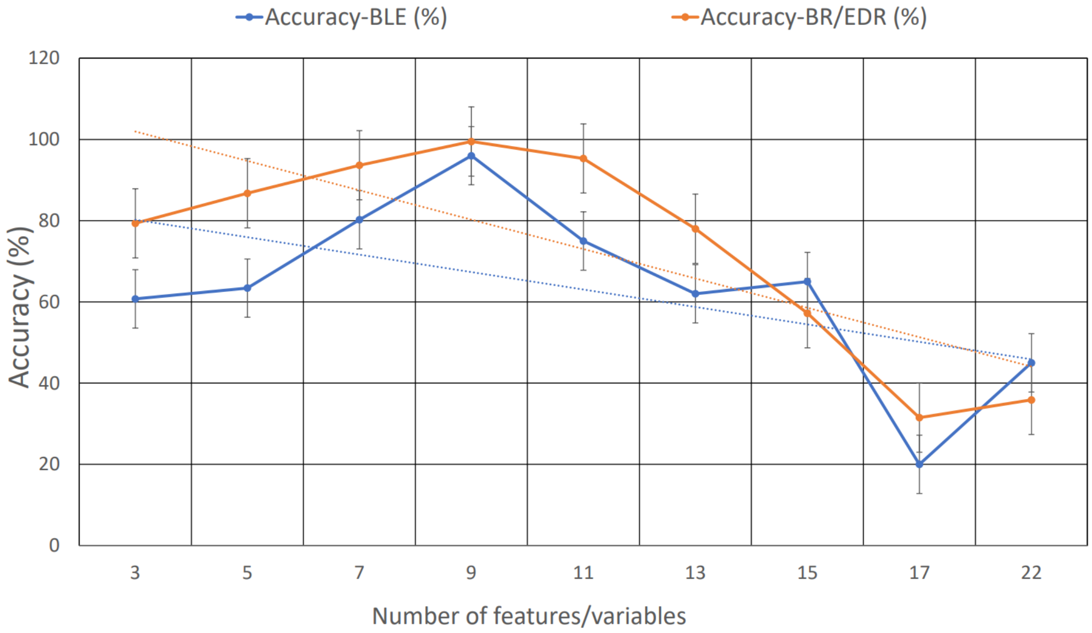

Figure 3.2: Accuracy of the model based on several features. *Based on the varying accuracy of the number of features, we chose 9 features from the dataset to train and test the model.*

We also utilized Logistic Regression (LR) [43] and Random Forest (RF) to determine which feature is contributing to the [44] output variable. Table 3.1 and 3.2 shows (”True” value) which indicates that the feature is contributing to the output variable, based on each univariate selection algorithm. The final score is given based on the cumulative of the four algorithms used. In the BR/EDR and BLE dataset, they contain 4 and 5 non-numerical values respectively. The non-numerical values are converted to numeric values before they are fed to the model using one-hot encoders, a process called numericalization. Finally, we select only the features that are important for identifying abnormal activities.

Normalization: It is a feature engineering technique used to have the data in one range for faster processing and classifier accuracy. There are various normalization techniques available, among which Z-score normalization is highly used due to its simplicity and performance accuracy [40].

Table 3.1: Univariate selection score of BR/EDR selected features

Feature	**Chi-2**	**RFE**	**LR**	**RF**	**Score**
btl2cap.length	True	True	True	True	4
HCI_events	True	True	True	True	4
HCI_ACL	True	True	True	True	4
Command Complete	True	True	True	True	4
Received direction	True	True	True	False	3
Sent Direction	True	True	False	True	3
Frame.cap_len	True	True	True	False	3
Disconnect complete	True	True	False	True	3
L2CAP	True	True	True	False	3

Table 3.2: Univariate selection score of BLE selected features

Feature	**Chi-2**	**RFE**	**LR**	**RF**	**Score**
btl2cap.length	True	True	True	True	4
Time	True	True	True	True	4
Protocol	True	True	True	True	4
Advertising	True	True	True	True	4
btle.AD	True	True	True	True	4
PPI.DLT	True	True	True	False	3
btatt.OM	True	True	False	True	3
btatt.OC	True	True	False	True	3

*Layer*_2

Initially, the medical data from IoT devices is collected and pre-processed on the first layer, and the collected events from Layer _1 events are sent for detection and identification to the second layer (the edge node). If any manipulation or deviation in the Bluetooth traffic is identified, an alert is triggered. On this layer, the events of the IoT medical device are actively captured and recorded on the events collector and are placed on the EPCRD device. This traffic is fed in the format of a feature vector which is represented in Eqn. 3.2

$$X(t) = (E_1, E_2, E_3, ..., E_n) \tag{3.2}$$

This feature vector is fed to Layer_2 to identify the malicious activities on this device based on the DL technique, which is deployed on the second layer of the edge node. The reason for placing two layers of intrusion detection is to protect the IoT system from device-based attacks and also to have full coverage of the IoT healthcare network. The classifier model gives 99% accuracy which has been placed on the Layer_2. As the pre-processing and intrusion detection phases are separated on different devices, the resulting system constitutes a multi-layer IDS. At last, the IDS model triggers an alert for the administrator to take the required course of action against the intrusion.

Dataset Description

We developed a Bluetooth (BR/EDR and BLE) dataset using realistic traffic generated using the smart healthcare testbed [45] described above Figure3.1. The dataset comprises abstract meta-information from the traffic flow of the Bluetooth-enabled medical IoT network. We have collected 5 GB of BR/EDR and BLE data over about 76 hours during the normal traffic patterns and also while performing the attacks. Therefore, the data collected is a combination of benign and malicious traffic. The performed attacks are DDoS, Bluesmack, MITM, and DoS on the L2CAP (Link Layer Control Adaption protocol) layer of the Bluetooth protocol stack. The L2CAP protocol is located in the data link layer of the stack, and it provides connection-less and connection-oriented data services to the top layer protocols. It allows the upper-level protocols and applications to send and receive the data frames. After analyzing the captured traffic in the pre-processing phase, we selected 9 features from each dataset through statistical methods and correlation analysis as presented in Tables 3.1 and 3.2.

IDS Classifiers

The entire classification process is divided into two main stages, training and testing. In the training phase, some samples of a dataset are used to train the model. In the testing phase, new samples are fed to the classifier from the test dataset to evaluate the performance. To validate the dataset performance, we used existing supervised and unsupervised ML algorithms in addition to the proposed DL model for training and testing. The reason for using various ML and the proposed DL models is to benchmark it and to show that the dataset is free from abnormal results on different classifier models. Many of the datasets used in the literature are algorithms dependent [46]. Our dataset produced acceptable accuracy for supervised and unsupervised ML and DL models. Various experiments with different classifiers helped us to build the most efficient DL model to identify malicious activities with more than 99% accuracy.

Classifier using Supervised ML Algorithms

Among the existing supervised ML algorithms, we have selected the most popular ones, namely: Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Random Forest (RF). We provide a short description of the algorithms that we have used in experiments.

Classifier using Unsupervised ML Algorithms

The selected algorithms are Naive Bayes (NB), Isolation Forest (IF), K-Means (KM), and Local Outlier Factor (LOF). Unsupervised algorithms are trained without using the labels of the features in the dataset. IoMT devices operate on different protocols, and due to this complexity, vulnerabilities may emerge. Furthermore, with classical ML algorithms, many attacks cannot be detected when the attacker does a small manipulation over time. DL techniques can recognize unknown patterns, outliers, and small changes from the training model.

Classifier using DNN

We used the Multilayer Perceptron (MLP) model, which is one of the categories of Feed-Forward Neural Network (FNN), with multiple layers: one input layer, one output layer, and three hidden layers. Each of these layers consists of a set of neurons. The process of assembling the hidden layers is known as a DNN as depicted in Figure 3.3. The DNN-IDS training comprises two phases - forward propagation and backward propagation. In forward propagation, output values are calculated. Whereas, in the backward propagation the weights are updated by passing the residual. The training of the model is implemented using Keras (with Tensorflow backend) and Table 3.3 provides

Table 3.3: DNN architectural hyper-parameters.

Description	**Setting**
Hidden Layer	3 (50, 25, 25)
Function	ReLU
Regularization	L2, dropout
Epochs	1000
Loss function	Binary_crossentropy
Optimizer	Adam
Batch Size	42
Dropout rate	0.25

detailed information on the various functions and parameters used. The combination of all layers is reflected in Figure 3.3. The model's hidden layers are formulated as in the MLP. The vector and the biases are represented as b_h and b_y.

$$f(\theta) = \mathbf{L}(y_t : \hat{y_t}) \tag{3.3}$$

- Hidden layer:

$$\textbf{\textit{Hl(x)}} = Hl_1(Hl_1 - 1(Hl - 2(....(Hl1(x))))) \tag{3.4}$$

- Training samples:

$$x_t = x_1, x_2, x_3, x_4, ..., x_{i-1}, x_i \tag{3.5}$$

- Hidden states:

$$h_t = h_1, h_2, h_3, h_4, ..., h_{i-1}, h_i \tag{3.6}$$

- Predictions of sequence:

$$\hat{y_t} = y_1, y_2, y_3, y_4, ...y_{i-1}, y_i \tag{3.7}$$

- Input-hidden weighted matrix:

$$Wl_x.Wl_h \tag{3.8}$$

- Output-hidden weighted matrix:

$$Wl_y \tag{3.9}$$

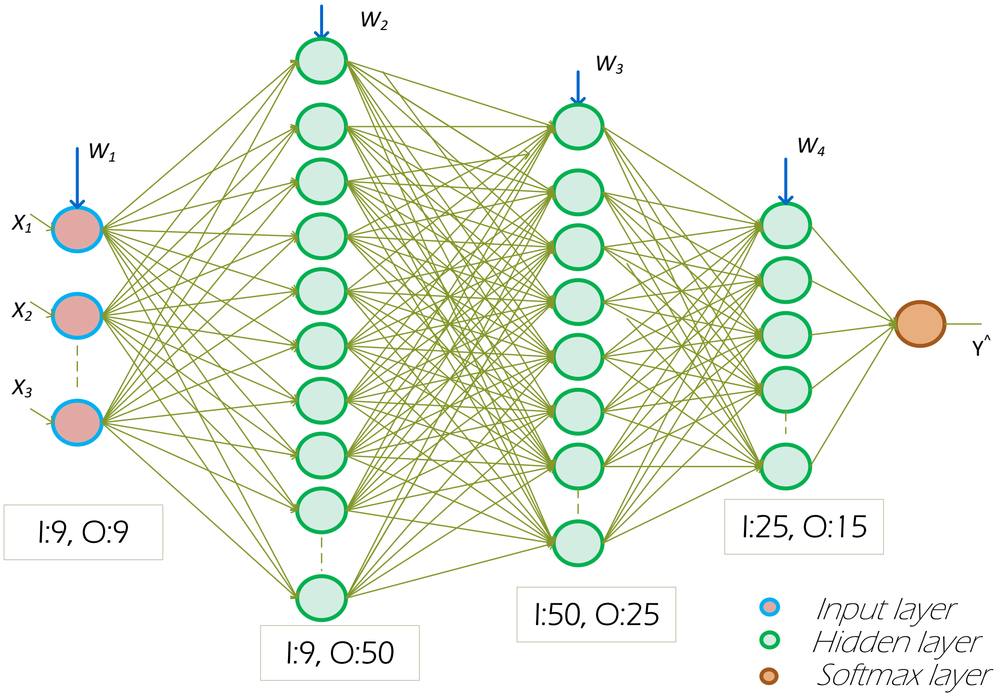

Figure 3.3: DNN architecture for the proposed IDS. It has three hidden layers with softmax as the output layer.

The objective function of the model, defined as the single pair of the training example (x_t, y_t) is:

L is described as the distance calculating the actual y_t and $\hat{y}_t$ denote the prediction labels, η denotes the learning rate and k denotes the number of iterations. In DNN, each hidden layer uses a non-linear activation function to model the gradient error. Among various activation functions, *ReLU* gives faster performance and can train the model with a huge number of hidden layers. For maximizing the efficiency of the

DNN, we build the model by considering the binary-cross entropy loss function, *ReLU* function, and *softmax* function with non-linear activation to achieve greater accuracy among the most substantial probability value of each class. In addition, we applied dropout techniques, to counter the problem of overfitting, by ignoring the randomly selected neurons. During this process, downstream neurons are ignored in the forward propagation and updated weights are not applied for the backward pass [47]. The neuron weights are settled within the network and are tuned for specific features. This effect on the network will result in less sensitivity to the definite weights of the neurons, which makes better generalization and is less likely to overfit the training data. In the below subsections, we show the experiments that we have performed in the selection of IDS classifiers for the IDS models.

CHAPTER 4: EXPERIMENTAL RESULTS

To choose the best classifier for Intrusion detection, we have trained and tested the BR/EDR and BLE Bluetooth datasets with supervised and unsupervised ML algorithms and DNN. The experimental results and discussion are provided below.

Un-Supervised ML Algorithms

BR/EDR Dataset

The BR/EDR dataset is trained and tested on 4 unsupervised ML algorithms with a balanced ratio of DOS attack and normal traffic pattern. We trained the 4 algorithms as binary classifiers to identify the DOS attack and normal traffic. The results achieved are shown in Table 4.1. Naïve Bayes algorithm recorded the highest accuracy, precision, F1-Score, and other favorable metrics among all the algorithms. The Precision and recall scores of Isolation Forest achieved an acceptable level of prediction, while K-means and LOF achieve more than 55% and 30% of precision and recall, respectively. This suggests that these two algorithms are not suitable to train the IDS using the created BR/ EDR dataset. Also, the reason for lower precision and recall of LOF is a direct indication that the dataset is fully pre-processed. The dataset does not contain a high level of deviations and we have performed intensive pre-processing on the dataset to make it normalized and free from outliers (in the Layer_1 of IDS model). Furthermore, the features that have been selected are highly significant for the output class. The other three metrics are F1 score, Area-under the ROC Curve (AUC), and Cohen's Kappa scores. These metrics provide a homogeneous pattern to the previous three metrics for the Naïve Bayes classifier.

Table 4.1: Performance analysis of the BR/EDR IDS using Unsupervised - ML algorithms

Metric	Naïve Bayes	Isolation forest	K-Means	LOF
Accuracy (%)	92.4	82.667	78.87	77.67
F1-score (%)	77.15	58.2	59.39	21.9
Recall (%)	63.68	52.34	55.01	38
Precision (%)	97.8	80.9	63.07	30.99
AUC (%)	82	59.38	53.48	51.62
Cohen's Kappa (%)	72.86	54.34	25.87	15.2

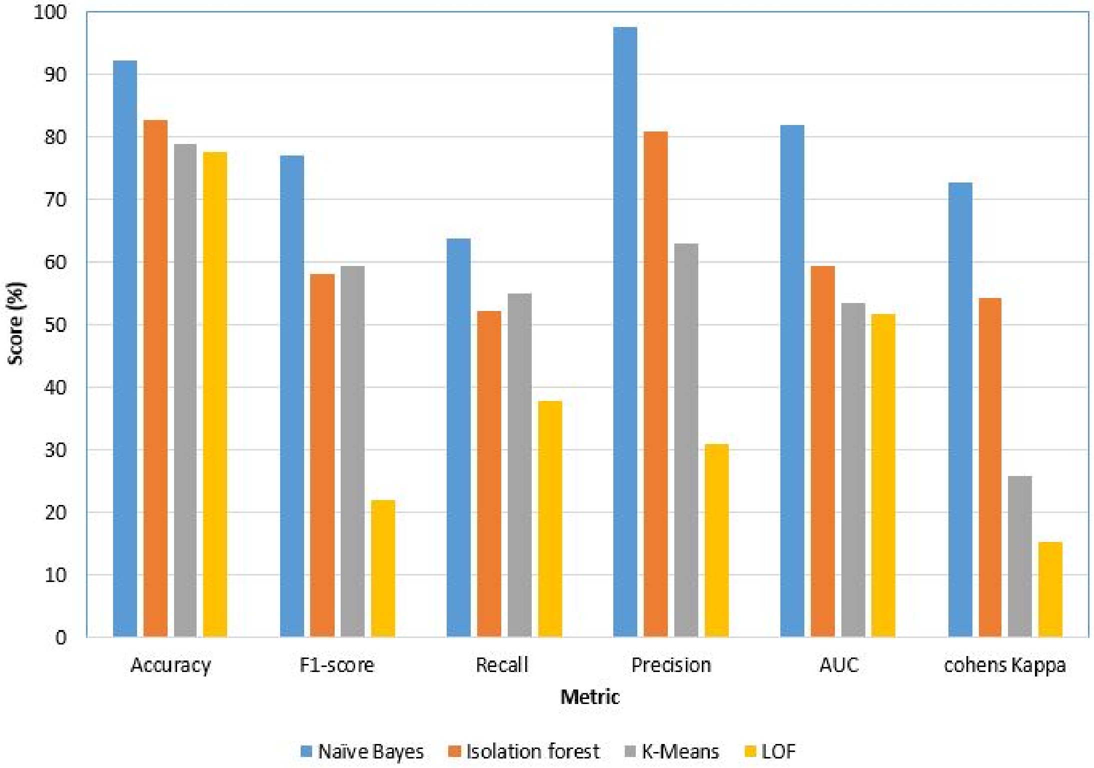

Figure 4.1: Performance of BR/EDR- Unsupervised ML algorithms. *This result shows that the dataset does not show any deviation irrespective of different models (i.e., the dataset is preprocessed intensively).*

Table 4.2: Performance analysis of multiclass classification of the BLE IDS using Unsupervised - ML algorithms. **1=DoS, 2=MITM, 3=Normal**

	Naïve Bayes (%)			Isolation forest (%)			K-means (%)			LOF		
Attacks ↪	1	2	3	1	2	3	1	2	3	1	2	3
Accuracy	98.7	78	80.4	79.4	70.7	87.09	80.2	74.2	88.23	81	67.4	70.7
F1-score	97.5	67	88	57.5	53.5	70.7	60.2	51.1	75	61.4	21.9	21.9
Recall	96.7	95	99	49.2	43.1	63.3	57	47.7	65.7	68	38	38
Precision	98.2	75	93	76.09	70.6	80.9	65	63.07	87.2	55.9	30.9	30.9
AUC	97.5	76	80	57.3	55.6	73.9	77.8	72.1	79.4	77.1	52.6	57.6
Cohen's Kappa	96	75.3	79.3	53.5	53.9	72.3	57.2	69.0	78.2	75	35.2	15.2

BLE Dataset

Similarly, the BLE dataset is trained and tested on the same unsupervised algorithms, but we modeled those as multiclass classifiers to identify DoS, MITM, and normal traffic from the samples. The performance of the classifiers is shown in Figure 4.2. The numeric scores of each class are visible from Table 4.2. Among the 4 unsupervised algorithms, Naive Bayes records the highest accuracy scores of $98, 78$ and 80 for DoS, MITM, and Normal traffic identification respectively. Recall, precision, and other metrics fall close to the accuracy scores for the Naive Bayes classifier. Isolation forest, K-means, and LOF classifiers show better performance than the BR/EDR dataset with an average accuracy of 80% for 3 classes.

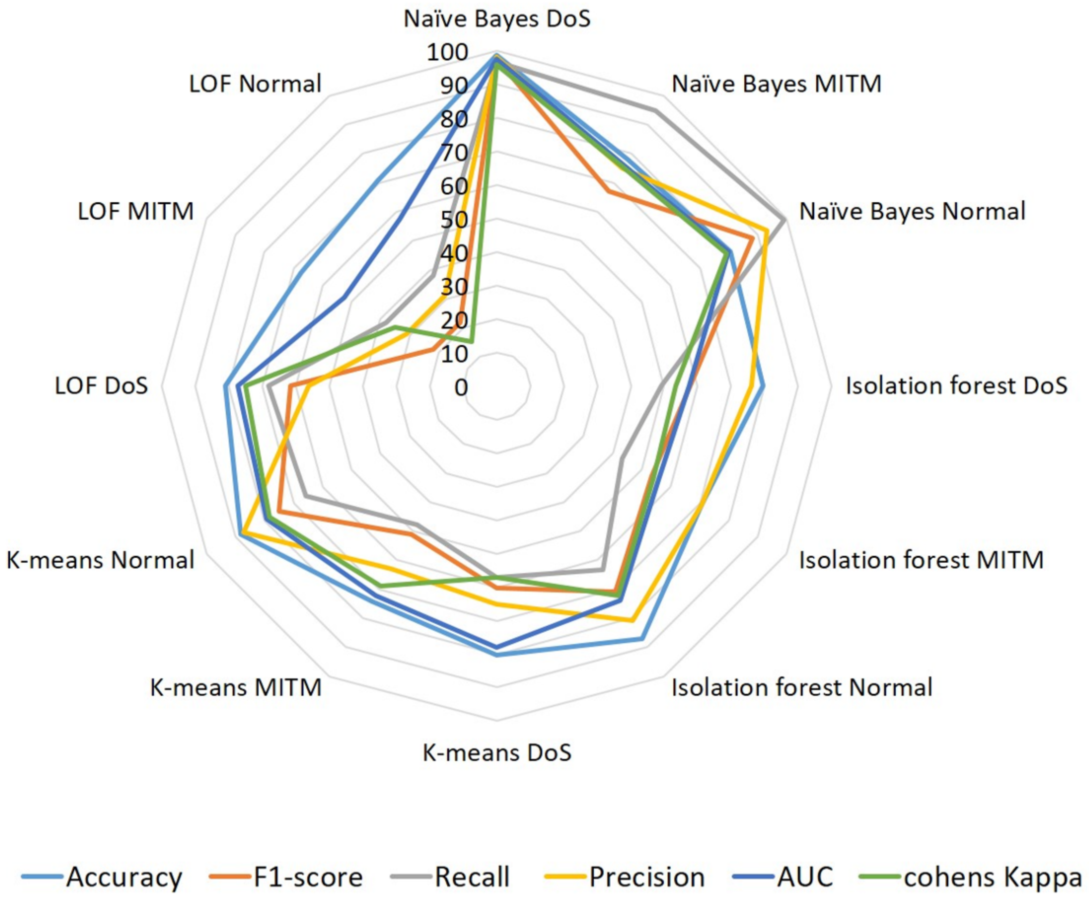

Figure 4.2: Performance of BLE- Unsupervised ML algorithms. *Multiple attacks are trained on the same models of BR/EDR, we observe that the models are not biased.)*

Supervised ML Algorithms

BR/EDR Dataset

Likewise, the dataset BR/ EDR is modeled as a binary classifier using four supervised ML algorithms each time, namely LR, DT, SVM, and RF to differentiate DoS attack and normal traffic. The experimental results depicted in Figure 4.3 and Table 4.3 show that Accuracy, Precision, and Recall are satisfactory for all classifiers. However, the RF classifier gives the highest score for all the 3 metrics, followed by DT, SVM, and then LR. This is clear evidence that the classifier model and dataset are efficient in identifying malicious traffic of DoS attacks on Bluetooth medical IoT devices.

Table 4.3: Performance analysis of the BR/EDR IDS using supervised - ML algorithms

	LR	DT	SVM	RF
Accuracy (%)	96.8	98.85	97.89	99.15
F1-score (%)	91.7	98.59	97.8	99.6
Recall (%)	88.32	98.5	96.6	98.6
Precision (%)	95.8	99.7	99.1	99
AUC (%)	94	100	98	100
Cohen's Kappa (%)	89.7	98.56	95.79	99.5

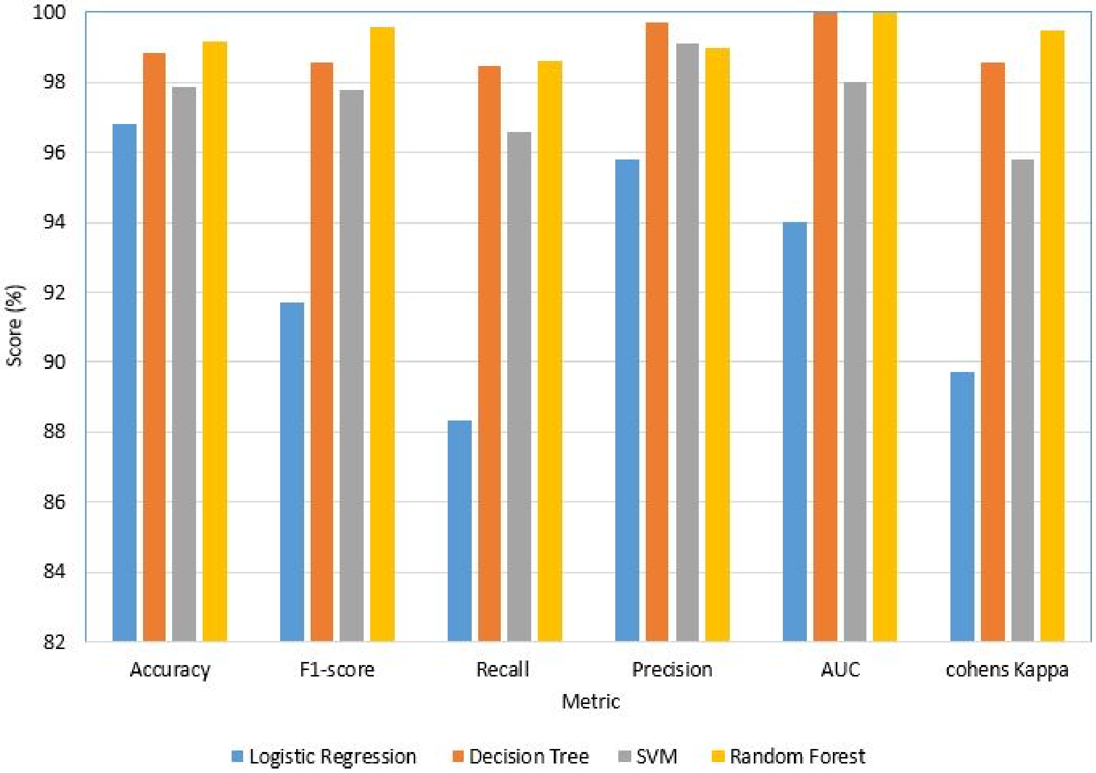

Figure 4.3: Performance of BR/EDR - Supervised ML algorithms. *The dataset and models are efficient in identifying malicious traffic behavior. (Deployed models are SVM and K-means.)*

Figure 4.3 also records the F1-Score, AUC score, and Cohen's Kappa score, which substantiates the inference that we deduced from the previous three metrics. Also, we can conclude that the dataset gives stable results using any of these supervised ML algorithms, of which RF and DT are the most recommended for general IoT devices and other networks. But, in the case of medical IoT devices, we need to choose a lightweight computationally inexpensive model. Among the tested algorithms, K-means (unsupervised) and SVM (supervised) are lightweight but they are computationally expensive in terms of training a model that is deployable on medical IoT devices. Nevertheless, the performance scores fall short for the real-time IDS model, so we investigate the DNN models using the created datasets.

BLE Dataset

The results of the multi-class model trained using BLE dataset with 4 different algorithms is shown in Figure 4.4 and Table 4.4. We observe that, unlike LR, the accuracy scores of the three supervised algorithms, DT, SVM, and RF lie between 95% and 98%. Though the average performance of the three algorithms, namely, DoS, MITM, and Normal, is satisfactory, it is difficult to choose the best among these three. Also, neither one of the single classifiers gives better performances for three classes identification to suit the real-time IDS performance. LR records less than 50% accuracy and unstable scores for other metrics. Because of these shortcomings, we investigate the use of a DNN model for both of the datasets.

Table 4.4: Performance analysis of multiclass classification of the BLE IDS using Supervised - ML algorithms. **1=DoS, 2=MITM, 3=Normal**

	LR (%)			DT (%)			SVM (%)			RF (%)		
Attacks ↪	1	2	3	1	2	3	1	2	3	1	2	3
Accuracy	48	79	94	96.63	98.5	97.29	97.89	94.39	96.86	97.74	96.5	95.78
F1 score	37	67	92	96.27	99.12	97.8	96.8	95	92	97.27	96.12	95.66
Recall	23	95	98	96.3	98.23	95.6	95.7	84	98	97.3	95.56	93.45
Precision	100	79	95	97.5	98.43	98.1	93.1	89	95	98.5	94.7	96.23
AUC	45	80	98	98	98.65	98	98	93	96	99	97.8	96.88
Cohen's Kappa	40	72	95	97	97.4	95.37	95.79	91.43	94.55	98	94	94.25

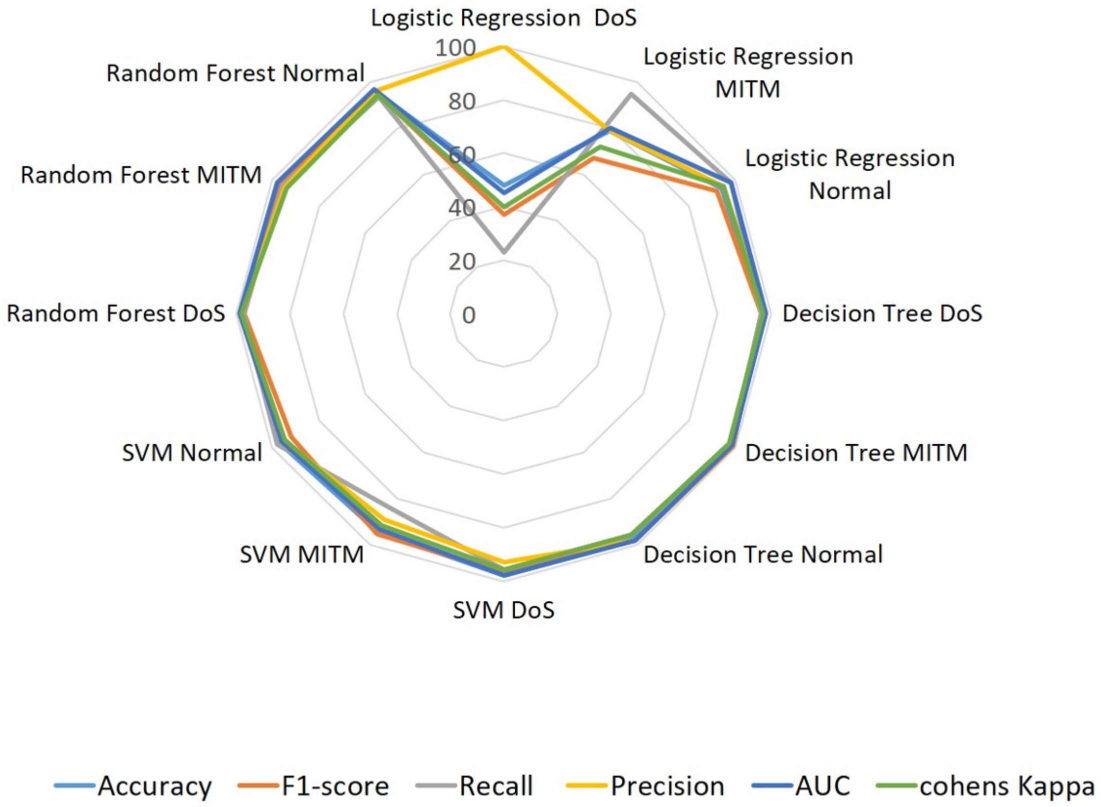

Figure 4.4: Performance of BLE- Supervised ML algorithms.*For real-time detection and deployment, neither of the single classifiers gives better performance.*

DNN model

Two DNNs are modeled as binary and multi-class classifiers using BR/EDR and BLE datasets respectively. The training accuracy of the two models lies between 92% and 95% as depicted in Figure 4.5. The testing accuracies are 98% and above for both models. From these results, we conclude that the classifier model using DNN is the best among all the other algorithms we tested. This deduction is bolstered by considering the Training and Testing Loss scores in Figure 4.6. The training loss of two models starts at approximately 0.3 and then reaches 0.15 as the learning process goes on. Similarly, the lowest Test loss recorded is 0.01 which is an indication of a stable DNN model.

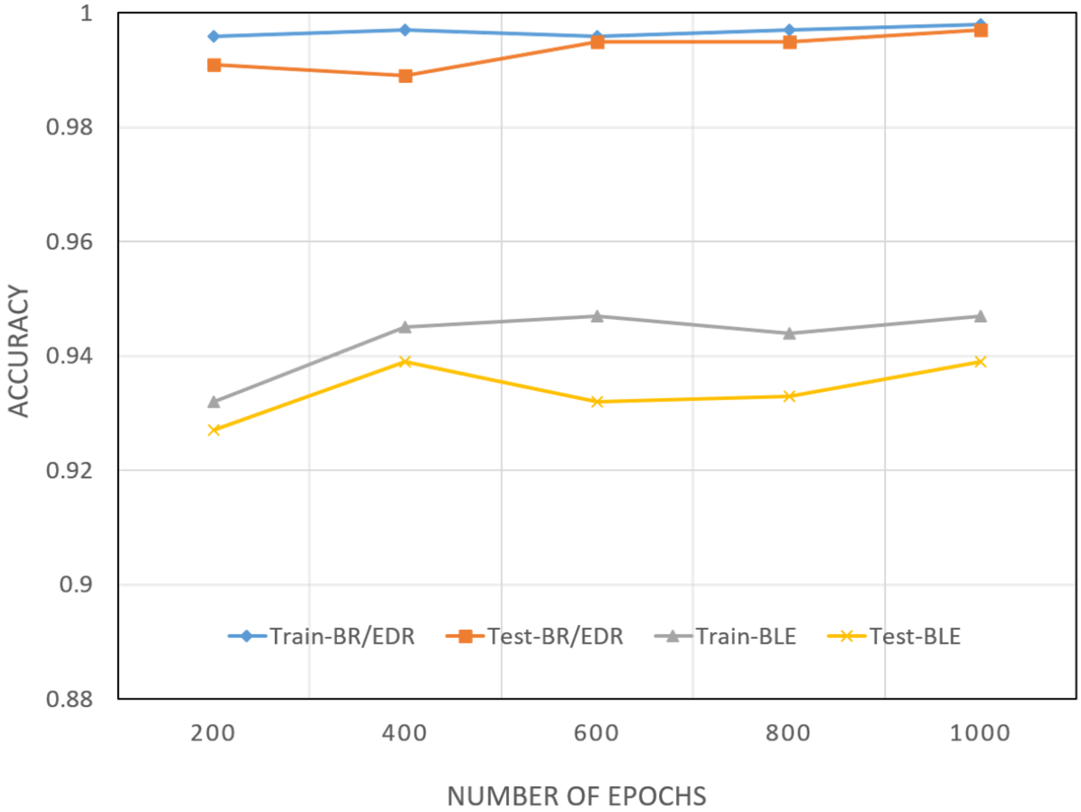

Figure 4.5: Training and Testing Accuracy. *The proposed IDS DNN model for BR/EDR and BLE dataset for 1000 epochs attains an accuracy of 98%.*

Additionally, to check the uniformity of the dataset, we have tested various ratios of abnormal (malicious) and benign traffic patterns. The ratios of benign and abnormal patterns considered are, $50-50$, $75-25$ and $80-20$. Each time the results that we achieved are consistent, which suggests that our dataset does not have any bias in the ratios of the traffic patterns. The accuracy scores of all the tests show that our dataset achieved less accuracy for unsupervised ML algorithms than for the supervised ML algorithms. From Table 4.5 and Figure 4.7, we deduce that the dataset can be considered a standard for training IDS models to identify DoS, DDoS, and Bluesmack attacks against Bluetooth IoMT devices. Moreover, in comparison to other models our proposed model attain best accuracy which is shown in Table 4.6.

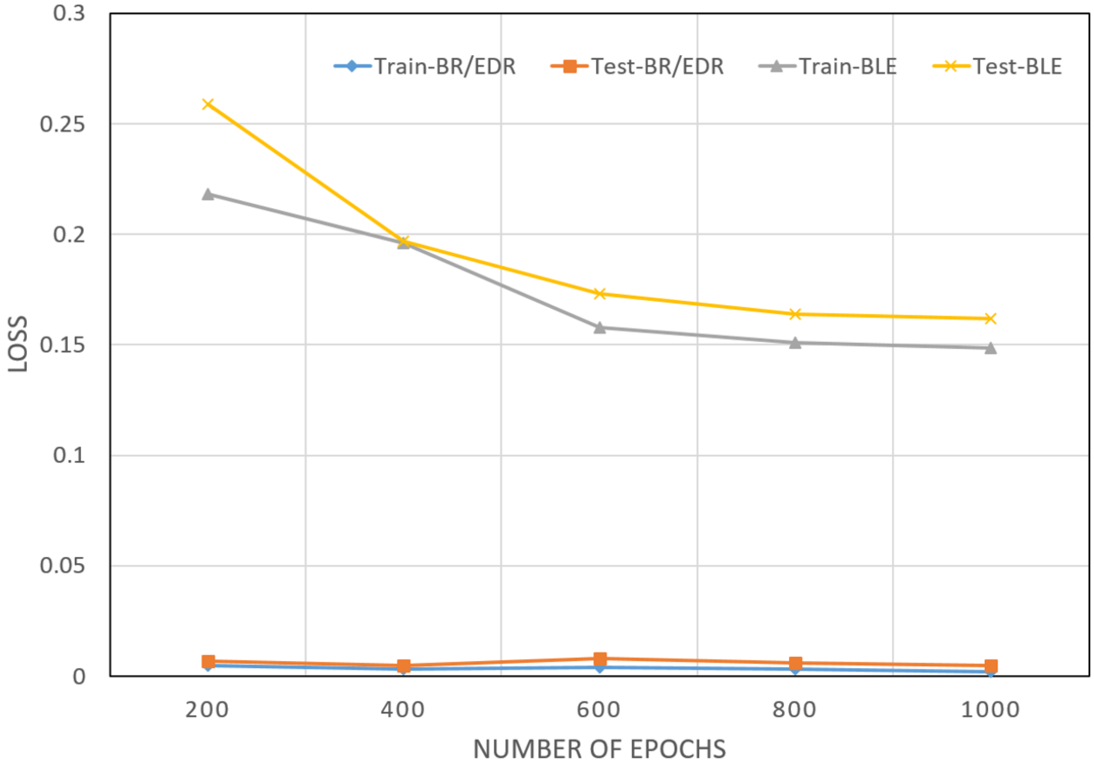

Figure 4.6: Training and Testing Loss - DNN. *The recorded test is loss of 0.01 that indicate DNN is reliable for real-time application,*

Table 4.5: Performance analysis of binary and multi-class classification of the proposed IDS (BR/EDR and BLE)

	BR/EDR	BLE-IDS			
	Binary	Binary	Multi class		
			DoS	MITM	Normal
Accuracy (%)	99.7	94.3	96.86	88.23	96.8
F1-score (%)	99.23	95	92	75	91.7
Recall (%)	98.65	84	98	65.78	88.32
Precision (%)	99.88	89	95	87.23	95.8
AUC (%)	99	93	96	79.43	94
Cohen's Kappa (%)	99.08	91.43	94.55	78.21	89.7

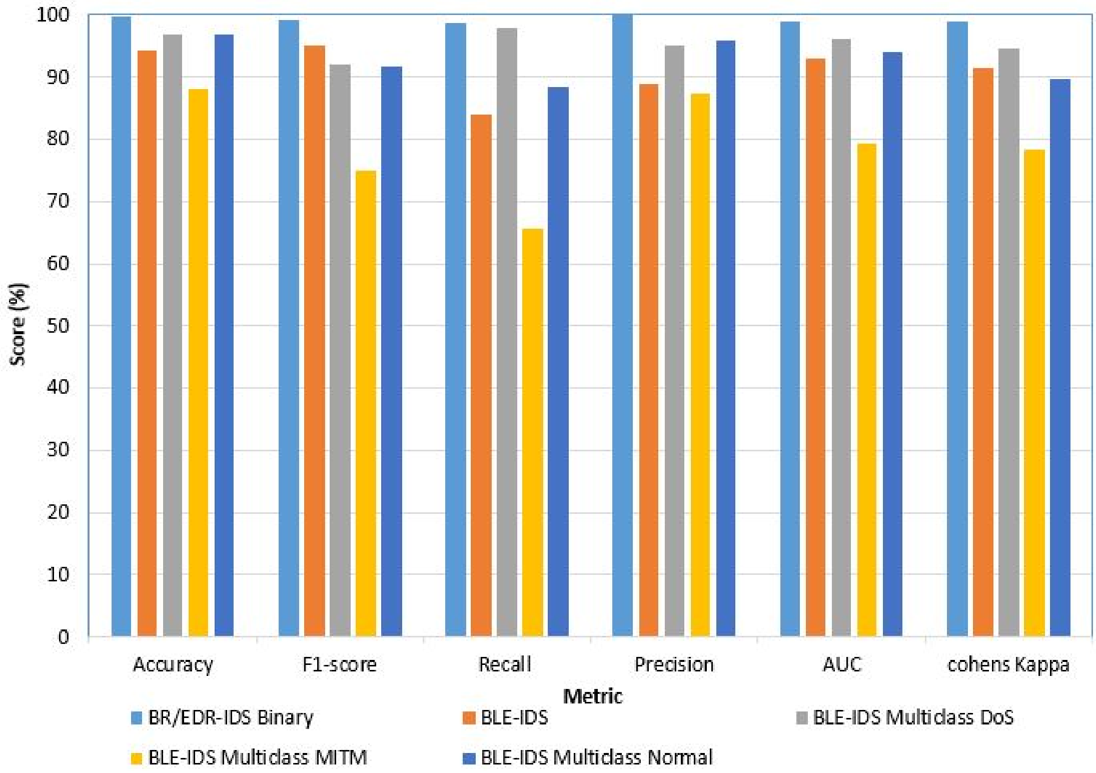

Figure 4.7: Performance analysis of binary and multiclass of the proposed model for BR/EDR and BLE respectively.

Table 4.6: Comparison of our model with existing IDS models

Model	Precision(%)	Recall(%)	F1(%)	Accuracy(%)
[38] (Bluetooth)	98	98	97	98.4
[35] (Bluetooth)	96.7	88.23	91.8	97
[48] (Bluetooth)	88.64	88.64	87.5	**
[49] (NSL-KDD)	95.72	98.65	**	97.06
[50] (NSL-KDD)	96	98.7	97.3	**
[51] (NSL-KDD)	**	98.6	**	99
Proposed IDS (BR/EDR)	99.7	99.06	99.38	**99.8**
Proposed IDS (BLE)	95	98	95	**96.86**

CHAPTER 5: CONCLUSIONS AND FUTURE WORK

Bluetooth communication is widely adopted in IoMT devices due to its various benefits. Nevertheless, because of its simplicity as a personal wireless communication protocol, Bluetooth lacks the security mechanisms which may result in devastating outcomes for patients treated using wireless medical devices. As discussed, continuous monitoring of network activity proves efficient in identifying cyber-attacks in most scenarios. We apply the same concept to Bluetooth-based medical IoT devices in a smart healthcare system. In this paper, we have proposed a secure and scalable architecture and deployed the IDS on the edge nodes of the smart healthcare system. The second outcome of this research is a standard Bluetooth dataset and a DNN-based classifier for Bluetooth traffic. To the best of our knowledge, this is the first intrusion detection dataset for the Bluetooth classic and BLE. From the results, we have seen that the created dataset can be used to train the IDS model for identifying DoS, DDoS, and Bluesmack attacks on medical IoT devices operating using Bluetooth technology. We also deduce that the proposed IDS classifier using DNN gives more than 99% accuracy, precision and recall, which outperforms the existing models for identifying Bluetooth-based attacks.

In the future, we plan to enhance the following critical areas of the proposed model. (1) We look forward to enlarging our dataset with more attack types, other than DoS, DDoS, and MITM. (2) We plan to include the attack data of other protocols such as WiFi. (3) After which, we aim to improve the intrusion detection classifier to identify those attacks efficiently on different datasets by applying data fusion or feature fusion techniques. (4) Furthermore, to develop a mitigation technique for the identified attacks from our model and also to detect unknown attacks so that the architecture can be extended to include mitigation mechanisms of the identified attacks.

REFERENCES

[1] "Internet of things (iot) connected devices installed base worldwide from 2015 to 2025 (in billions),"

[2] P. K. Khatua, V. K. Ramachandaramurthy, P. Kasinathan, J. Y. Yong, J. Pasupuleti, and A. Rajagopalan, "Application and assessment of internet of things toward the sustainability of energy systems: Challenges and issues," *Sustainable Cities and Society*, vol. 53, p. 101 957, 2020.

[3] D. Das and J. Zhang, "Pandemic in a smart city: Singapore's covid-19 management through technology & society," *Urban Geography*, vol. 42, no. 3, pp. 408–416, 2021.

[4] I. Alrashdi, A. Alqazzaz, E. Aloufi, R. Alharthi, M. Zohdy, and H. Ming, "Ad-iot: Anomaly detection of iot cyberattacks in smart city using machine learning," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, 2019, pp. 0305–0310.

[5] ISACA, "The ultimate list of healthcare it statistics for 2020," `https://arkenea.com/healthcare-statistics,2020`, 2020.

[6] R. O. Andrade, S. G. Yoo, L. Tello-Oquendo, and I. Ortiz-Garcés, "A comprehensive study of the iot cybersecurity in smart cities," *IEEE Access*, vol. 8, pp. 228 922–228 941, 2020.

[7] A. Limaye and T. Adegbija, "A workload characterization for the internet of medical things (iomt)," in *2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, IEEE, 2017, pp. 302–307.

[8] Online, “83% of medical devices run on outdated operating systems,” `https://www.hipaajournal.com/83-of-medical-devices-run-on-outdated-operating-systems`, 2020.

[9] A. M. Lonzetta, P. Cope, J. Campbell, B. J. Mohd, and T. Hayajneh, “Security vulnerabilities in bluetooth technology as used in iot,” *Journal of Sensor and Actuator Networks*, vol. 7, no. 3, p. 28, 2018.

[10] N. A. Gunathilake, A. Al-Dubai, and W. J. Buchana, “Recent advances and trends in lightweight cryptography for iot security,” in *2020 16th International Conference on Network and Service Management (CNSM)*, IEEE, 2020, pp. 1–5.

[11] D. Popoviæ, D. Bojaniæ, N. Jorgovanoviæ, S. Dosen, and R. Petroviæ, “Teleecg based on bluetooth transceivers,” in *Telecommunications Forum TELFOR*, vol. 20, 2003.

[12] U. M. Rijah, S. Mosharani, S. Amuthapriya, M. Mufthas, M. Hezretov, and D. Dhammearatchi, “Bluetooth security analysis and solution,” *International Journal of Scientific and Research Publications*, vol. 6, no. 4, pp. 333–338, 2016.

[13] N. Park, B. K. Mandal, and Y.-H. Park, “Sensor protocol for roaming bluetooth multiagent systems,” *International Journal of Distributed Sensor Networks*, vol. 9, no. 4, p. 963 508, 2013.

[14] B. A. Miller and C. Bisdikian, *Bluetooth revealed: the insider’s guide to an open specification for global wireless communication*. Prentice Hall PTR, 2001.

[15] Y. Liu, S. Li, and L. Cao, “Application of bluetooth communication in digital photo frame,” in *2009 ISECS International Colloquium on Computing, Communication, Control, and Management*, IEEE, vol. 4, 2009, pp. 370–373.

[16] C. Bisdikian *et al.*, “An overview of the bluetooth wireless technology,” *IEEE Commun Mag*, vol. 39, no. 12, pp. 86–94, 2001.

[17] J. C. Haartsen, “Bluetooth radio system,” *Wiley Encyclopedia of Telecommunications*, 2003.

[18] J. Haartsen, M. Naghshineh, J. Inouye, O. J. Joeressen, and W. Allen, “Bluetooth: Vision, goals, and architecture,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 2, no. 4, pp. 38–45, 1998.

[19] A. Palin, J. Salokannel, and J. Reunamaki, *Method and system for establishing a wireless communications link*, US Patent 7,352,998, Apr. 2008.

[20] R. U. Rasool, H. F. Ahmad, W. Rafique, A. Qayyum, and J. Qadir, “Security and privacy of internet of medical things: A contemporary review in the age of surveillance, botnets, and adversarial ml,” *Journal of Network and Computer Applications*, p. 103 332, 2022.

[21] Y. Yuehong, Y. Zeng, X. Chen, and Y. Fan, “The internet of things in healthcare: An overview,” *Journal of Industrial Information Integration*, vol. 1, pp. 3–13, 2016.

[22] M. Zubair, A. Ghubaish, D. Unal, *et al.*, “Secure bluetooth communication in smart healthcare systems: A novel community dataset and intrusion detection system,” *Sensors*, vol. 22, no. 21, p. 8280, 2022.

[23] P. Cope, J. Campbell, and T. Hayajneh, “An investigation of bluetooth security vulnerabilities,” in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, 2017, pp. 1–7.

[24] V. P. Musale and S. Apte, “Security risks in bluetooth devices,” *International Journal of Computer Applications*, vol. 51, no. 1, 2012.

[25] Online, “Sweyntooth vulnerabilities in ble chips affect many medical devices,” `https://www.hipaajournal.com/sweyntooth-vulnerabilities-in-bluetooth-low-energy-chips-affect-many-medical-devices`, 2020.

[26] G. Franzè, D. Famularo, W. Lucia, and F. Tedesco, “A resilient control strategy for cyber-physical systems subject to denial of service attacks: A leader-follower set-theoretic approach,” *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1204–1214, 2020.

[27] H. Zhang, Y. Li, Z. Lv, A. K. Sangaiah, and T. Huang, “A real-time and ubiquitous network attack detection based on deep belief network and support vector machine,” *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 3, pp. 790–799, 2020.

[28] F. Salo, A. B. Nassif, and A. Essex, “Dimensionality reduction with ig-pca and ensemble classifier for network intrusion detection,” *Computer Networks*, vol. 148, pp. 164–175, 2019.

[29] S. P. RM, P. K. R. Maddikunta, M. Parimala, *et al.*, “An effective feature engineering for dnn using hybrid pca-gwo for intrusion detection in iomt architecture,” *Computer Communications*, vol. 160, pp. 139–149, 2020.

[30] A. A. V. Rani and E. Baburaj, “Secure and intelligent architecture for cloud-based healthcare applications in wireless body sensor networks,” *International Journal of Biomedical Engineering and Technology*, vol. 29, no. 2, pp. 186–199, 2019.

[31] E. Anthi, L. Williams, M. S lowińska, 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.

[32] V. Gazis, "A survey of standards for machine-to-machine and the internet of things," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 482–511, 2016.

[33] A. Tabassum, A. Erbad, and M. Guizani, "A survey on recent approaches in intrusion detection system in iots," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, IEEE, 2019, pp. 1190–1197.

[34] K. M. Haataja, "New efficient intrusion detection and prevention system for bluetooth networks," in *1st International ICST Conference on Mobile Wireless Middleware, Operating Systems and Applications*, 2010.

[35] M. Krzysztoń and M. Marks, "Simulation of watchdog placement for cooperative anomaly detection in bluetooth mesh intrusion detection system," *Simulation Modelling Practice and Theory*, vol. 101, p. 102 041, 2020.

[36] P. Satam, S. Satam, and S. Hariri, "Bluetooth intrusion detection system (bids)," in *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, IEEE, 2018, pp. 1–7.

[37] S. Satam, P. Satam, and S. Hariri, "Multi-level bluetooth intrusion detection system," in *2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)*, IEEE, 2020, pp. 1–8.

[38] A. I. Newaz, A. K. Sikder, L. Babun, and A. S. Uluagac, "Heka: A novel intrusion detection system for attacks to personal medical devices," in *2020 IEEE Conference on Communications and Network Security (CNS)*, IEEE, 2020, pp. 1–9.

[39] Online, "More than half of IoT devices vulnerable to severe attacks," `https://threatpost.com/half-iot-devices-vulnerable-severe-attacks/153609/`, 2021.

[40] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *Ieee Access*, vol. 5, pp. 21 954–21 961, 2017.

[41] S. Han, K. Zhu, M. Zhou, and X. Cai, "Information-utilization-method-assisted multimodal multiobjective optimization and application to credit card fraud detection," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 4, pp. 856–869, 2021.

[42] H. Liu, M. Zhou, and Q. Liu, "An embedded feature selection method for imbalanced data classification," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 3, pp. 703–715, 2019.

[43] P. Bolourchi, M. Moradi, H. Demirel, and S. Uysal, "Improved sar target recognition by selecting moment methods based on fisher score," *Signal, Image and Video Processing*, vol. 14, no. 1, pp. 39–47, 2020.

[44] I. S. Thaseen, C. Kumar, *et al.*, "Intrusion detection model using chi square feature selection and modified naıve bayes classifier," in *Proceedings of the 3rd International Symposium on Big Data and Cloud Computing Challenges (ISBCC–16')*, Springer, 2016, pp. 81–91.

[45] M. Zubair, D. Unal, A. Al-Ali, and A. Shikfa, "Exploiting bluetooth vulnerabilities in e-health iot devices," in *Proceedings of the 3rd international conference on future networks and distributed systems*, 2019, pp. 1–7.

[46] C. Liu, S. Chakraborty, and P. Mittal, "Dependence makes you vulnberable: Differential privacy under dependent tuples.," in *NDSS*, vol. 16, 2016, pp. 21–24.

[47] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[48] J. Roux, E. Alata, G. Auriol, M. Kaâniche, V. Nicomette, and R. Cayre, "Radiot: Radio communications intrusion detection for iot-a protocol independent approach," in *2018 IEEE 17th International Symposium on Network Computing and Applications (NCA)*, IEEE, 2018, pp. 1–8.

[49] J. Kim, J. Kim, H. Kim, *et al.*, "An approach to build an efficient intrusion detection classifier," *Journal of Platform Technology*, vol. 3, no. 4, pp. 43–52, 2015.

[50] G. Thamilarasu and S. Chawla, "Towards deep-learning-driven intrusion detection for the internet of things," *Sensors*, vol. 19, no. 9, p. 1977, 2019.

[51] A.-H. Muna, N. Moustafa, and E. Sitnikova, "Identification of malicious activities in industrial internet of things based on deep learning models," *Journal of information security and applications*, vol. 41, pp. 1–11, 2018.

PUBLICATIONS

- Zubair, M.; Ghubaish, A.;Unal, D; Al-Ali, A; Reimann, T; Alinier, G; Hammoudeh, M; Qadir, J.Secure Bluetooth Communication in Smart Healthcare Systems: A Novel Community Dataset and Intrusion Detection System. Sensors 2022, 22
- Zubair, Mohammed, et al. ”Exploiting bluetooth vulnerabilities in e-health IoT devices.” Proceedings of the 3rd international conference on future networks and distributed systems. 2019.
- Zubair M, Unal D, Al-Ali A, Reimann T, Alinier G. Cybersecurity for next generation healthcare in Qatar. Journal of Emergency Medicine, Trauma and Acute Care. 2021 Aug 9;2021(2-Qatar Health 2021 Conference abstracts):41.
- Aggarwal M, Zubair M, Unal D, Al-Ali A, Reimann T, Alinier G. Fuzzy Identification-Based Encryption for healthcare user face authentication. Journal of Emergency Medicine, Trauma and Acute Care. 2022 Jan 15;2022(1-Qatar Health 2022 Conference abstracts):72.
- Aggarwal M, Zubair M, Unal D, Al-Ali A, Reimann T, Alinier G. A Testbed Implementation of a Biometric Identity-Based Encryption for IoMT-enabled Healthcare System. InThe 5th International Conference on Future Networks & Distributed Systems 2021 Dec 15 (pp. 58-63).