A Study on Designing of Energy Efficient and Secure Data Communication System for IoT Wireless Network

A Study on Designing of Energy Efficient and Secure Data Communication System for IoT Wireless Network

IoT無線ネットワークの電力効率化・セキュア化を
実現する通信システムの設計に関する研究

A dissertation presented
by

Shikhar

submitted to
Tohoku University
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

Supervisor: Professor Nei Kato

Department of Applied Information Sciences
Graduate School of Information Sciences
Tohoku University

January, 2021

Abstract

With emergence of Internet of Things (IoT), new network challenges such as massive connection, energy efficiency, and security are arising. To deal these challenges, industries and academia have been proposing the next-generation wireless communication technologies for IoT. 5G has been proposed for wireless wide area network (WWAN) which includes IoT communication. Similarly, IEEE 802.11ah is designed for IoT communication in the wireless local area network(WLAN), also known as Wi-Fi HaLow. However, each technology has trade-off in terms of meeting the IoT network requirements. For instance, 5G, specially long term evolution-advanced (LTE-A) Pro network has included features such as group paging, Narrow-band IoT, end-to-end security and so on to enable massive and secure communication. However, 5G have not designed any specific energy efficient protocols and cellular networks consume higher energy compared to other wireless technologies. Therefore, in this thesis, IoT data communication technology of LTE-A Pro networks i.e.; group paging is improved by minimizing energy consumption while providing required quality of services. Likewise, IEEE 802.11ah has included novel medium access control protocol such as restricted access window (RAW), traffic indication map for energy efficient communication by massive amount of devices. However, IEEE 802.11ah has not provided any new security protocols for IoT communication, which can be an acute concern for network security. Therefore, this thesis investigates Internet-wide port scan approach over IEEE 802.11ah for risk assessment of IoT devices and proposes novel IEEE 802.11ah network-aware IWPS system to maximize the IoT security. Overall, this thesis proposes novel approaches for energy efficient and secure data transmission over the next-generation IoT wireless network.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The Internet of Things (IoT) is one of the emerging technologies that promises to transform our society into an “intelligent society” by enabling and connecting smart “sensor embedded physical objects” to the Internet [1]. With the interconnection of smart objects, IoT provides innovative services such as intelligent transport systems, smart healthcare and home, intelligent industry, and so on. Based on predictions by IHS Markit, the number of connected devices is going to be massive, i.e., around 125 billion [2]. Moreover, IoT devices are miniature compared to traditional devices which make them constrained by energy and resources such as memory. Thus, advanced connectivity alternatives need to be developed for meeting IoT communication requirements at each level of the network. To address IoT connectivity necessities, standards such as the Institute of Electrical and Electronics Engineers (IEEE), and 3rd Generation Partnership Project (3GPP) have proposed various wireless communication technologies for each type of network [3, 4]. Wireless communication technologies are obligatory for IoT owing to the ubiquitous connection requirement by smart objects such

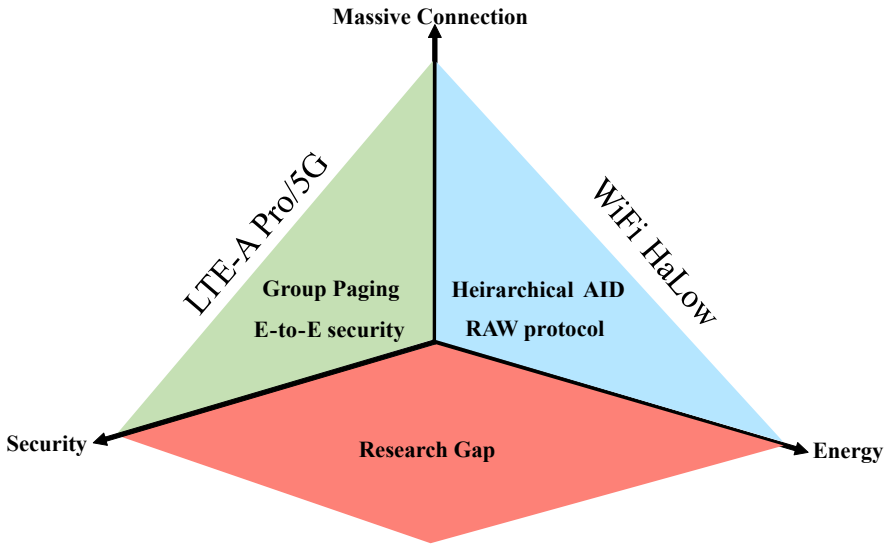

Figure 1.1: **Research Gap**: Energy-efficient and Secure IoT data communication over LTE-A Pro and IEEE 802.11ah respectively

as intelligent cars. Therefore, this thesis focuses on IoT wireless network. Wireless networks are usually classified based on scale and range of networks such as wireless local area network (WLAN), wireless wide area network (WWAN), and wireless personal area network (WPAN). Similarly, IoT wireless network is classified into these types and specific communication protocols are designed for each type. This thesis will emphasize on WLAN and WWAN IoT communication technologies. The latest cellular technology i.e., long term evolution advanced (LTE-A) Pro over 5G includes technologies for enabling IoT data communication over WAN. Thus, this thesis will focus on the IEEE 802.11ah and 5G communication technologies for IoT data communication.

Massive connection, low energy consumption and secure transmission are the major requirements for IoT data communication, as shown in the three axis of Fig. 1.1 [5]. IoT devices will be available everywhere owing to various innovative application, which massively deploy devices in every area. Thus, communication

technologies should support massive connection. Moreoever, IoT devices should run for decades with a single battery owing to their placement in sophisticated places, which makes replacement and recharging difficult. Moreover, the size of devices also restricts them from having high power batteries. Therefore, IoT data communication should be energy efficient. IoT security is another challenge due to constrained resources such as memory, processing, and so on; restricting the implementation of security protocols, intrusion protection systems, and complex passwords [7]. IoT devices are prone to exposure due to placement in open public places. Therefore, IoT devices are easy prey for attackers. Moreover, the infected devices can act as bots, that can be exploited to initiate distribute denial services (DDoS) attack on other network infrastructure [7]. IoT security becomes an area of acute concern which needs to be addressed for IoT data communication. Each IoT communication technology should be energy efficient and secure while enabling massive connections. However, both IEEE 802.11ah and 5G technologies have a trade-off in addressing the above issues.

5G technologies, especially long-term evolution-advanced (LTE-A) pro has proposed technologies for supporting massive communication such as group paging, as depicted in Fig. 1.1 [6]. Group paging is a novel pull-based data communication approach that extracts data from a group of devices periodically [8]. It avoids congestion at cellular networks. Moreover, cellular networks are more secure than Wi-Fi due to use of licensed spectrum, well-protected infrastructure under the control of mobile service providers, and end-to-end encryption with internet security protocols at the back end that provides protected communication over 5G, as mentioned in Fig. 1.1. However, high energy consumption by mobile devices is always a problem for cellular networks [8]. Therefore, energy-efficient design of IoT data communication system over LTE-A Pro networks such as group paging should be focused.

IEEE 802.11ah has also designed group based new medium access control (MAC) protocol including features such as restricted access window (RAW), traffic indication map, associated identification, and time to wake up, as mentioned inside the area under WiFi-Halow in the Fig. 1.1 [9]. IEEE 802.11ah supports massive connections and provide energy-efficient communication, as depicted in Fig. 1.1. However, IEEE 802.11ah has not proposed improved security features for novel IoT security issues, which inherits existing security issues [10]. Thus, this thesis focuses on improving IoT security over the IEEE 802.11ah. In this regard, researchers are investigating to use a well-known network sifting mechanism i.e.; Internet-wide port scan (IWPS) for identifying vulnerabilities in the IoT [11, 12]. However, existing scanners lack the probing with WLAN awareness [12]. Moreover, IEEE 802.11ah has low bandwidth compared to traditional IEEE 802.11n. Thus, improper scan rate without awareness of IEEE 802.11ah efficiency can lower the performance network, which may increase the risk on IoT devices. Therefore, network-aware IWPS system is needed for improving security over IEEE 802.11ah.

1.2 Research Purposes

This thesis aims to design models for energy efficient and secure IoT data communication over IoT wireless network. Especially, this thesis focuses on modeling group paging technologies over LTE-A pro networks and IWPS over IEEE 802.11ah network for improving IoT security. Therefore, this thesis proposes new models to achieve following purposes:

1. **Design grouping approaches for mobile IoT devices (MIDs) to minimize energy consumption while providing QoS**

2. **Design network-aware IWPS for IEEE 802.11ah enabled IoT devices to maximize IoT security**

For the first purpose, this thesis proposes mathematical models to estimate energy and QoS parameters such as delay and packet loss for grouping of MIDs having various traffic characteristics and mobility pattern. Thereafter, to achieve the first purpose, this thesis provides optimal solution such that energy consumption is minimized while satisfying the QoS constraints. The proposed models and solution enable energy-efficient IoT data communication using group paging approach over LTE-A Pro networks.

For the second purpose, this thesis proposes novel network-aware IWPS mathematical models to maximize the security of IoT devices and improve the performance of IWPS over IEEE 802.11ah. The proposed model considers "risk" as security metric to analyze the security of any IoT devices. These proposed models evaluates the impact of low IoT network performance on cyber security services such as confidentiality, integrity, availability of IoT data, and attack likelihood owing to IWPS. Based on these models, this thesis minimizes the risk on IoT devices by providing optimal scan rate for IWPS such that IEEE 802.11ah network has better performance to provide security services. Such proposed models can assist network operators and security admins to efficiently set scan rate for IWPS without compromising IoT security services and network's performance.

1.3 Thesis Structure

The remainder of this thesis is organized as follows.

The overview of data communication technologies over IoT wireless networks are presented in Chapter 2. Moreover, the related works of each focused technologies are pointed out in this chapter.

Chapter 3 describes research challenges of grouping approach of MIDs. Thereafter, this chapter proposes novel mathematical models to optimize the energy consumption and meet the QoS requirements of each IoT device. Finally, this chapter explains analysis of the proposed models while comparing its performance to random grouping approach.

Chapter 4 presents background on focused security technology for IoT and research challenges of IWPS related to IoT security degradation. To solve such challenges, this chapter proposes mathematical models to maximize IoT security based on scan rate and network performance of IEEE 802.11ah network. Finally, this chapter shows the numerical analysis to maximize the security.

Finally, Chapter 5 draws concluding remarks to this thesis.

Chapter 2

Data Communication System over IoT Wireless Networks

2.1 Introduction

This chapter introduces data communication technologies of IoT wireless network. As stated in the previous chapter that LTE-A Pro under 5G is the focused technology for communication over WWAN and IEEE 802.11ah for WLAN. Hence, at first, this chapter presents LTE-A Pro communication technologies such as paging, discontinuous reception (DRX), and group paging. Thereafter, this chapter presents the related works of group paging with the research gap in this technology. Moreover, for IEEE 802.11ah, this chapters explains it's novel medium access control (MAC) protocol that are restricted access window (RAW), and carrier sense multiple access/collision avoidance (CSMA/CA). This chapter focuses on security issues of IoT over WLAN, and introduces the security analysis technology i.e. IWPS. Thereafter, this chapter describes the related works of IWPS over IEEE 802.11ah based on performance of MAC protocol.

The content of this chapter are referred from these papers that are written

and published by me and my co-authors.

- Shikhar Verma, Yuichi Kawamoto and Nei. Kato, "Energy-Efficient Group Paging Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G," in *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 9187-9199, Oct. 2019.

- Shikhar Verma, Yuichi Kawamoto and Nei. Kato, "A Network-aware Internet-wide Scan for Security Maximization of IPv6-enabled WLAN IoT Devices," in *IEEE Internet of Things Journal* (Accepted).

- Shikhar Verma, Yuichi Kawamoto and Nei. Kato, "Security Analysis of Network-Oblivious Internet-Wide Scan for IEEE 802.11ah Enabled IoT," *IEEE International Conference on Vehicular Communications (VTC-Fall 2020), Virtual Conference,* Nov. 2020.

2.2 Data Transmission Technology over LTE-A Pro

2.2.1 Paging and DRX

LTE-A Pro networks has included several features for efficient data transmission by IoT devices such as paging and DRX mechanism [13]. Paging is a pull-based data transmission approach used by cellular networks. This method reduces network congestion by restricting a device to access the channel at an instant. In paging approach, mobile management entity transmits the PM to all eNodeB (eNBs) in the tracking area list, and then each eNB broadcasts the paging message (PM) at the Paging Occasion (PO) of that specific device, as depicted in Figure. 2.1. Devices monitor the physical downlink control channel for PM at each PO. Upon receiving the PM, a device initiates random access channel (RACH) process in the next random access slot. If a device successfully receives the uplink resources from the network during RACH procedure, then it starts the uplink data transmission. Figure. 2.1 shows the PO for Device 1 and Device 2 only. Although monitoring for PM at each radio sub-frames can be attractive from the delay aspect, it comes at the expense of high energy consumption of devices. Since IoT devices are restrained by power consumption and data transmission is periodic, IoT devices should follow the DRX mechanism of LTE networks after data transmission to reduce such energy consumption. DRX mechanism allows IoT devices to turn off their radio interface and enter into sleep mode after data transmission/reception [14]. Hence, every device has a predefined, periodic sleep-wakeup cycle, also called its DRX cycle. As shown in Figure. 2.1, devices wake up after sleep mode at their PO to check for any PM. If a device receives the PM, it initiates RACH procedure and uplink data transmission. A device can also download data from eNB after receiving PM. In case there is no PM, the device goes

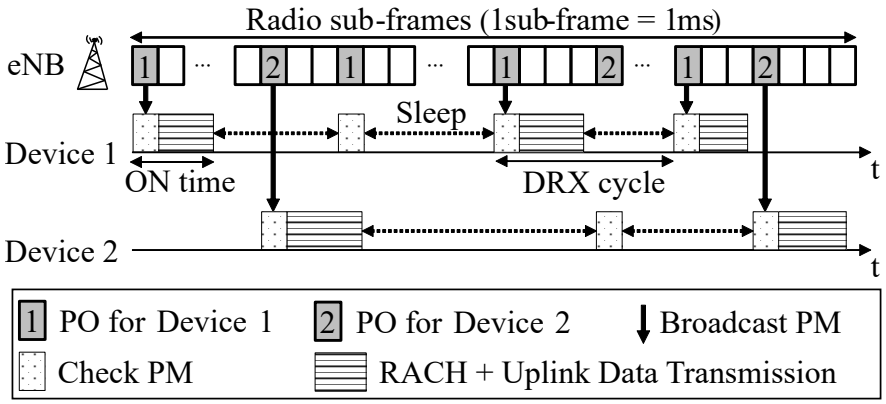

Figure 2.1: Paging and DRX mechanism of LTE-A Pro networks

to sleep mode without any transmission/reception. However, we are considering periodic uplink data transmission in IoT. Hence, we can say that each device has DTF. We define DTF as the number of times uplink data is transmitted within a particular period. For instance, in Figure. 2.1, Device 1 initiates uplink data transmission thrice, and Device 2 does so twice within a certain period of time. Hence, IoT devices have two characteristics, namely DRX cycle (sleep-wake) and DTF. In this paper, we consider these two characteristics of IoT devices. DRX mechanism in LTE-A Pro networks is known as extended-DRX because DRX cycle can be longer for IoT devices that need to sleep for hours. Hence, DRX and e-DRX has same meaning in this thesis. However, eNB has to broadcast a prohibitively large number of PMs to activate a large number of devices over many time frames; thus rendering paging impractical when the number of IoT devices is massive [8]. Moreover, multiple PMs create massive signal overload at the network. Therefore, 3GPP proposed a *group paging* mechanism to deal with the limitations of paging mechanism for IoT devices.

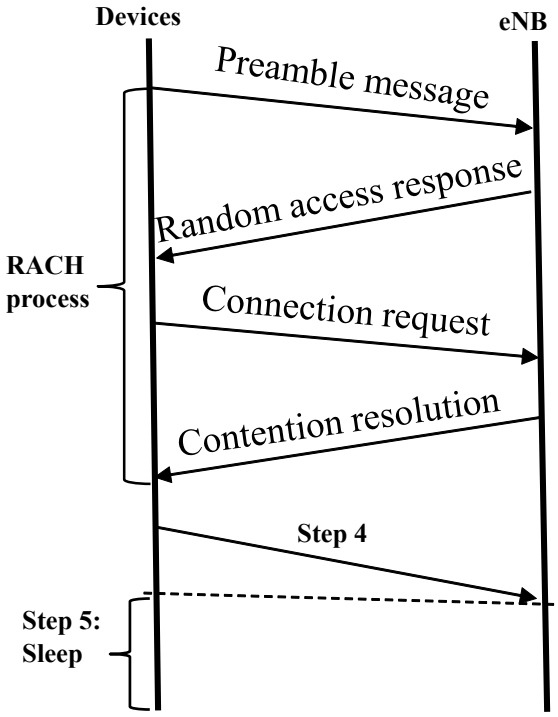

Figure 2.2: RACH process by devices of group

2.2.2 Group paging

The group paging is an efficient pull-based mechanism for periodic data transmission from a massive number of IoT devices that overcomes long paging delays by activating a group of devices through a single PM [8]. In group paging, there are five steps. In the first step, IoT devices are arranged into different groups, and eNB assigns a unique GID to each IoT device upon joining the cell and a group. The devices in a group are denoted as *aG-IoTDs* in this thesis. In step 2, eNB broadcasts the group ID (GID). In step 3, *aG-IoTDs* having same GID initiate the RACH process simultaneously at the first available random access slot after receiving PM. The RACH process is a four-step process to obtain uplink synchronization and data transmission resources, as shown in Figure. 2.2. In step 4, *aG-IoTDs* starts the data transmission after successfully establishing connection

to the eNB through RACH process [15]. After data transmission, *aG-IoTDs* go to sleep mode in step 5. As we can see, in the group paging, *aG-IoTDs* wake-up and initiate data transmission simultaneously . Hence, *aG-IoTDs* should have common DRX cycle and DTF. The explanation to set DRX and DTF of *aG-IoTDs* is discussed in the next subsection.

2.2.3 Group Paging Constraints

As mentioned in previous subsection, each IoT device has two required characteristics, namely, DRX and DTF, and group paging allows a group of devices to wake-up, receive PM and start the data transmissions simultaneously. Therefore, to synchronize the PM notification, sleep-wake cycle and uplink data transmission, all *aG-IoTDs* should have a common DRX and DTF that can be called as the group DRX (G-DRX) and group DTF (G-DTF). As a result, group paging changes the required characteristics of *aG-IoTDs*. Therefore, G-DRX and G-DTF of a group should be determined such that devices in a group meet their required DRX and DTF. Hence, G-DRX should be set such that each *aG-IoTD* should have at most their original DRX, and G-DTF should be selected such that each *aG-IoTD* can transmit at least equal to its original DTF. Thus, G-DRX is chosen as minimum DRX of *aG-IoTDs* in the group, and G-DTF is the maximum DTF of *aG-IoTDs* in the group. Hence, the network decides the G-DRX and G-DTF of each group based on the required characteristics of *aG-IoTDs*. These restrictions are referred to as group paging constraints in the rest of the paper. The group paging constraint is represented from (2.1) to (2.2):

$$GDX_G = \min(DX_1^G, DX_2^G, \ldots, DX_N^G), \tag{2.1}$$

$$GF_G = \max(F_1^G, F_2^G, \ldots, F_N^G). \tag{2.2}$$

GDX_G and GF_G are the G-DRX and G-DTF of group G, respectively. DX_i^G and F_i^G are the required or original DRX and DTF of *aG-IoTD* i of group G, respectively. N is the total number of *aG-IoTDs* in a group.

Fig. 2.3 depicts an example of group paging constraints. In Fig. 2.3, an IoT Device, named Device 1, of a smart logistic truck and another IoT Device, named Device 2, of a smart car are group together. Hence, Device 1 and Device 2 are *aG-IoTDs* after grouping. Device 1 and Device 2 have DRX1 and DRX2 as their original DRX cycles, where DRX1 is shorter than DRX2. Similarly, Device 1 and Device 2 have two and three times required uplink data transmission within a certain time, respectively. After grouping, G-DRX is configured to the DRX of Device 1 to meet the group paging constraints, i.e. (2.1). The DRX of Device 2 modifies to the DRX of Device 1, which means more wake up within a certain period of time. Similarly, from (2.2), G-DTF is set to three. Hence, DTF of Device 1 changes to three times within a certain time. Finally, it can be concluded

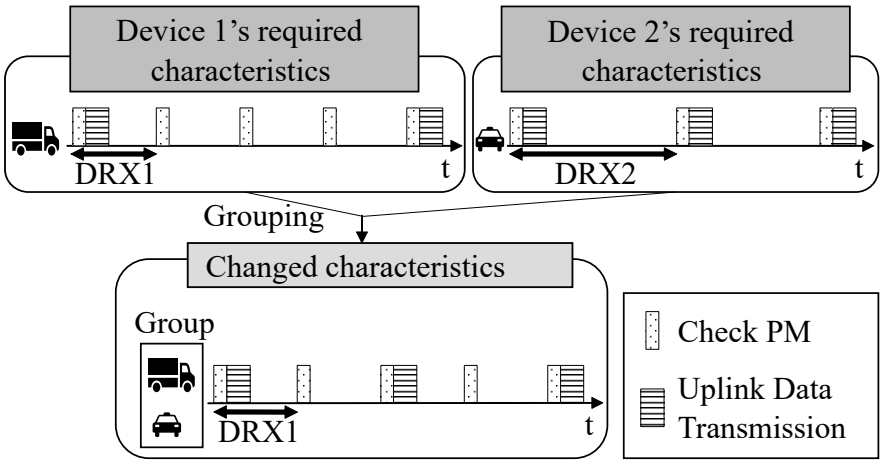

Figure 2.3: An illustration of group paging constraints.

that owing to group paging constraints, DRX and DTF of group depends on types of *aG-IoTDs* and *aG-IoTDs* are decided based on the grouping method. This chapter presents the related work to the group paging technology in the next section.

2.2.4 Related Works to Group Paging

This section discuses the existing works for each step of group paging approach. There are several works addressing methods to design step 2, 3, and 4. For instance,for the step 2, [20] proposed approaches to efficiently schedule the group paging message such as consecutive group paging. Similarly,for step 3 and 4, [8, 15, 16] addressed several issues such as improvement of the RACH regarding preamble collision probability, access success probability, dynamic resource allocation, access delay and so on. There are also existing works to optimize DRX for the step 5 of group paging approach. Existing literature on DRX has studied performance improvement of DRX mechanism by tuning the parameters such as ON time, inactivity timers, sleep ratio and so on, which optimize the energy consumption based on traffic characteristics and QoS features [14, 17, 18]. However, most researchers have considered optimizing DRX of a device in paging mechanism without considering the performance improvement of DRX of devices in group paging mechanism, which may increase the energy consumption and degrade QoS. Hence, this chapter does not focus on performance improvement of DRX of individual IoT devices; instead, this chapter focuses on DRX of groups based on devices in the group. However, researchers including the 3GPP standard have not extensively studied grouping methods for IoT devices. As, we have seen that characteristics of a group such as DRX and DTF depends on the devices in the group. Moreover, these characteristics are key parameters which can impact energy consumption and QoS of any IoT devices in the group. Hence, this the-

sis addresses the grouping challenge of IoT devices in group paging concerning energy consumption and QoS.

2.3 Data Transmission Technology over IEEE 802.11ah

IEEE 802.11ah is a WLAN protocol, designed to meet wide array of IoT uses cases and its requirements. This protocol is used by Wi-Fi alliance and also called as Wi-Fi HaLow. It is build upon the success of IEEE 802.11n, operating in sub-gigahertz frequency band. Enable to work in such lower bands provides several benefits such as long range (around 1Km), penetration through obstacles (low propagation loss), Internet Protocol (IP) connectivity, and so on. This protocol also provide variable data rates (150 kbps to 100 Mbps), which is required for heterogeneous IoT applications. Hence, IEEE 802.11ah is well suited technology for IoT with IP connectivity, better data rate and coverage than other low power IoT wireless network protocols such as Sigfox, Wi-SUN etc. The IEEE standard has introduced several novel medium access (MAC) protocols features to support massive connection under an AP and save significant amount of energy compared to other WLAN standards. Hence, in this section, we presents the MAC protocols of IEEE 802.11ah and its research gap i.e; security issues. Finally, this section discuses the approaches for improvement of IoT security over IEEE 802.11ah networks with their related works.

2.3.1 MAC protocol of IEEE 802.11ah

IEEE 802.11ah has proposed restricted access window (RAW) protocol for the channel access by massive amount of IoT devices [9]. This protocol divides the time such as beacon interval into several RAW frame and each RAW frame

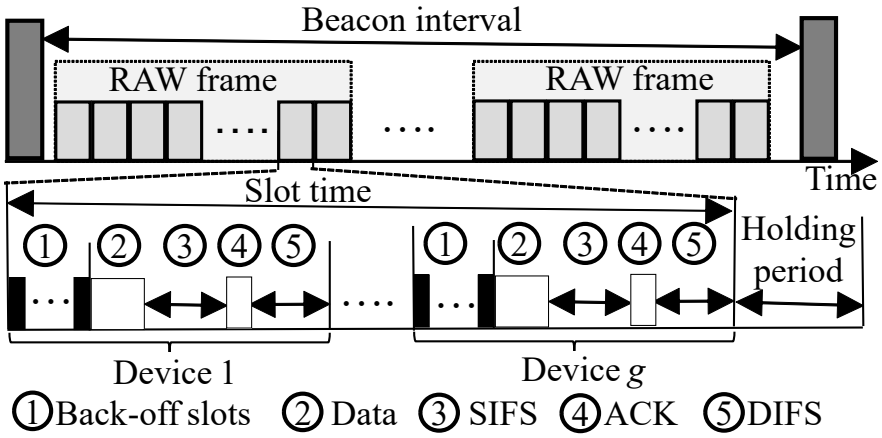

Figure 2.4: MAC Protocol of IEEE 802.11ah

(RFrame) is slotted into fixed number of RAW slots (Rslot). Each slot is dedicated for a group of IoT devices for channel access and data transmission. Hence, the IoT devices under an AP are divided into several groups. This protocol reduces channel contention and saves significant energy consumption by allowing devices to sleep in other than their specified Rslots. Moreover, devices in a group follow Carrier-sense multiple access with collision avoidance (CSMA/CA) approach for sub-GHz channel access, as shown in Figure. 2.4. In Figure. 2.4, we can see that g number of devices of a group access the channel during Rslot using CSMA/CA processs. At first, each device of a group starts the back-off slots. After the end of backoff-slots, device starts the data transmission if it senses ideal channel, otherwise device initiate another random backoff. In case device finds idea channel, that device initiate the data transmission followed by short inter frame space (SIFS) time to receive acknowledgment (ACK) from AP. After receiving ACK, channel is idle for distributed interframe space (DIFS) time to allow other devices to sense idle channel. A Rslot is handover to another group after completion of CSMA/CA process by group. However, there is an issue during handover of Rslot between two groups that is whether to allow the ongoing transmission is

allowed to it's Rslot boundary or not. IEEE 802.11ah defines "No-Crossing slot" if it is not allowed. IEEE 802.11ah includes holding period to avoid the crossing of transmission to another slot. Devices are not allowed to transmit if they can't initiate before holding period, as shown in Figure. 2.4. In case of allowing to cross slot, IEEE 802.11ah defines this case as "Crossing slot". However, no new transmission is allowed after completion of ongoing transmission in case of crossing slot. This thesis considers non-crossing slot case. However, IoT security over IEEE 802.11ah is still concern. Therefore, next section presents the security issues of IoT data transmission over IEEE 802.11ah.

2.3.2 Security Challenges of IoT over IEEE 802.11ah

IEEE 802.11ah has not proposed any new security features considering IoT's constraint like memory, processing and so on. The upgrade of Wi-Fi protected access (WPA) is included in the IEEE 802.11ah. However, devices will experiences similar problem as traditional AP. Moreover, WPA security methods may not be able to implement in constrained IoT devices. Moreover, IoT devices are publicly available which means it is easy to physically access and tamper with its security. IoT devices have weak password also and unnecessary ports which welcome the vulnerabilities. The patching management system is also not well defined owing to massive IoT devices, placed in sophisticated places. Intrusion protection system (IPS) is also weak due to low processing resources which restricts the deployment IPS. Hence, security of IoT components needs to be monitored regularly. In this regard, researchers and industries are going to use vulnerability scan at each device through regular port scan. Hence, this thesis focuses on the regular port scan of IoT devices to identify the vulnerability and improve security. In this thesis, our target is to scan Internet-wide connected IoT devices. Hence, we consider the Internet-wide port scan (IWPS) of IoT devices.

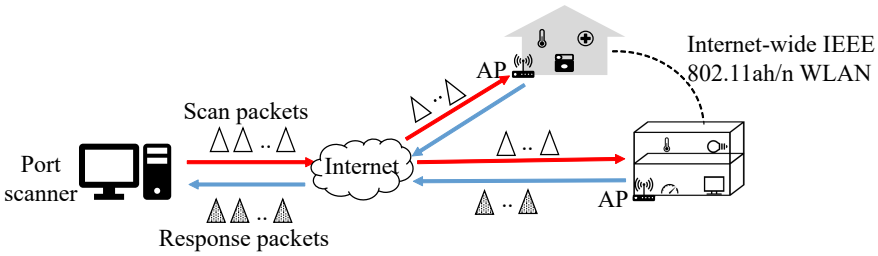

Figure 2.5: Internet-wide port scan approach

2.3.3 Internet-wide port scan (IWPS)

The IWPS is an approach to scan services running on each port of the Internet-wide connected devices, as shown in Figure. 2.5. The objective of IWPS is to identify the any vulnerabilities in the Internet-wide connected devices. There are various port scan techniques such as open, half-open, stealth scan and so on. In this study, we consider transmission control protocol (TCP) half-open scan owing to low burden on network and devices. With the TCP half-open scan, the scanner generates synchronization (SYN) packets at a certain scan rate. The SYN packets are scan packets in the Figure. 2.5. With a TCP SYN request, an IoT device can respond with a reset (RST) in the case of closed ports or an acknowledgment (ACK) in the case of open ports, as shown in Fig. 2.5. The ACK and RST packets are response packets in the Figure. 2.5. The port scanner can RST the ports after receiving an ACK. Thereafter, the scanners analyze the responses to detect vulnerabilities. They then notify the telecommunication carriers to alert the device owners and support centers, which, in turn, suggest patching solutions. The scan and response packets has to transmit through the IEEE 802.11ah networks. Hence, the performance of port scan depends on the IEEE 802.11ah network performance and vice versa.

2.3.4 Related Works to IWPS

Vulnerability scanning was developed for local private network-scanning capabilities for less intrusive and IWPS for common public IPs. Most studies on the reconnaissance of connected device ports were related to the design of efficient frameworks/scanners using different approaches to generate and distribute SPs for IWPSs [21, 22]. Other studies focused on intrusion detection and prevention by analyzing the scan responses [23, 24]. However, most of these studies did not address the WLAN network limitations. For example, the ZMAP product claims that it is capable of scanning an entire IPV4 address in under 45 min for a given port [21] using 97% of the gigabit ethernet capacity, which is not practical for IoT, owing to the unavailability of such throughput, plus the congestion caused by Het-IoT traffic. The Masscan product faces the same issue. However, both ZMap and Masscan have overcome the long scan delays issue of tools, such as network mappers, by generating SPs faster and distributing them randomly. However, the random distribution of SPs can overload IoT networks at certain access points (AP) where fewer network resources are available for ultra-dense devices [25]. In [25], a model was proposed to set a high scan rate while maximizing the throughput for traditional IEEE 802.11 in order to deal with network constraints. However, this model considered only network improvements and ignored security, including the impact on IPsec services, which is a prime objective of the port scan. Moreover, the model was suited only to traditional IEEE 802.11; it was not applicable to the IEEE 802.11ah RAW mechanism. Hence, in this study, we address the high scan rate, which is not necessarily beneficial in terms of IoT security. Moreover, emerging energy-efficient WLAN MAC protocols, such as IEEE 802.11ah RAW, can restrict the scanning of sleeping IoT devices. Hence, disorganized and random probing of IoT device ports over IEEE802.11ah can degrade network and port-scan performance. Furthermore, most existing stud-

ies on IEEE 802.11ah have focused on grouping strategies and performance of the RAW protocol only, whereas the performance of traditional scan traffic under such IoT-centric protocols has not been investigated [9, 26]. Therefore, this study investigates a pioneering network-aware IWPS for IPV6-enabled IoT device security over the latest IEEE802.11ah networks. Moreover, we probe devices on specific APs together rather than randomly to avoid congestion at uncertain APs. In the next section, we discuss the research challenges faced by legacy port scanning over IEEE 802.11ah.

2.4 Summary

This chapter presents prominent data communication technologies of LTE-A Pro networks and IEEE 802.11ah WLAN. Group paging approach is explained in this chapter, which is an effective data transmission technology for massive IoT connection. Moreover, this chapter provides the related works of group paging. Based on related works of group paging, this chapter presents a research issue to investigate grouping approach of IoT devices concerning energy and QoS. Furthermore, this chapter explains the MAC protocol of IEEE 802.11ah and security challenges over it. In order to improve IoT security, this chapter presents a key technologies i.e. IWPS and its related work. However, IWPS needs to be also investigated owing to network-oblivious scan, which can degrade the performance.

Chapter 3

An Energy Efficient Grouping Approach for QoS constrained mobile IoT devices

3.1 Introduction

This chapter presents the research challenge concerning inappropriate grouping of MIDs with *aG-IoTDs* in a cell, which leads to significant energy consumption. Moreover, this chapter investigates the scenarios where MIDs can have long packet arrival delay (PAD) and packet loss rate (PLR) due to inappropriate grouping. To address such issues, this chapter proposes the novel mathematical models to minimize energy consumption while meeting QoS requirements. Finally, this chapter provides mathematical analyses and performance evaluation of the proposed approach.

The parts of contents in this chapter are referred to the following papers that are written based on our own researches.

- S. Verma, Y. Kawamoto and N. Kato, "Energy-Efficient Group Paging

Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G," in IEEE Internet of Things Journal, vol. 6, no. 5, pp. 9187-9199, Oct. 2019.

3.2 Research Challenges: Energy and QoS

As in the previous chapter, we need to investigate the grouping techniques of MIDs based on diverse characteristics and mobility patterns. Hence, this section discusses system model and research challenges faced by MIDs upon joining a new group.

3.2.1 High Energy Consumption

In this research, we consider that K groups are available in a cell, each containing N *aG-IoTDs*, and MIDs join a new group G upon changing its current cell, where $G \in \{1, 2, 3, \ldots, K\}$. Thus, there is a continual change in the configuration of groups due to joining and disbanding of groups by MIDs. Moreover, from (2.1) to (2.2), we can see that *aG-IoTDs* have two common characteristics, namely GDX_G and GF_G owing to the group paging constraint. Each MID has its own required DRX and DTF, which should meet after joining a group. To meet the constraints, grouping of MIDs can change the group's characteristics or their own characteristics. Hence, there can be two scenarios where the grouping of MIDs can change the characteristics of *aG-IoTDs* and MIDs.

The first scenario happens when DRX of new MIDs is greater than the G-DRX or vice-versa. The DRX of MIDs is changed to G-DRX in case of greater DRX based on (2.1). Hence, MIDs wake up more frequently leading to high energy consumption. Similarly, G-DRX can be greater than MIDs' DRX. In this case, the *aG-IoTDs* change their DRX to MIDs' DRX, which increases their energy

consumption. The second scenario happens when there is a difference in the DTF of new MIDs and G-DTF. The DTF of MIDs becomes more frequent if it is less than G-DTF as expressed in (2.2). Hence, MIDs have more frequent data transmission, thus increasing the energy consumption of MIDs. Furthermore, the DTF of MIDs can be greater than G-DTF. In this case, the DTF of *aG-IoTDs* is changed to MID's DTF which entails more frequent data transmissions for those *aG-IoTDs* and resulting in increased energy consumption. Therefore, the major difference between DRX of new MIDs and *aG-IoTDs* or DTF of them may lead to high energy consumption. It may be possible that inappropriate grouping of such diverse characteristics increase energy consumption significantly. Therefore, our objective in this paper is to find optimal grouping for MIDs such that the total energy consumption is minimum.

3.2.2 QoS degradation

In our system model, we consider MIDs having different mobility patterns. MIDs move from a cell to another, as shown in Fig. 3.1(a). A person with a wearable IoT

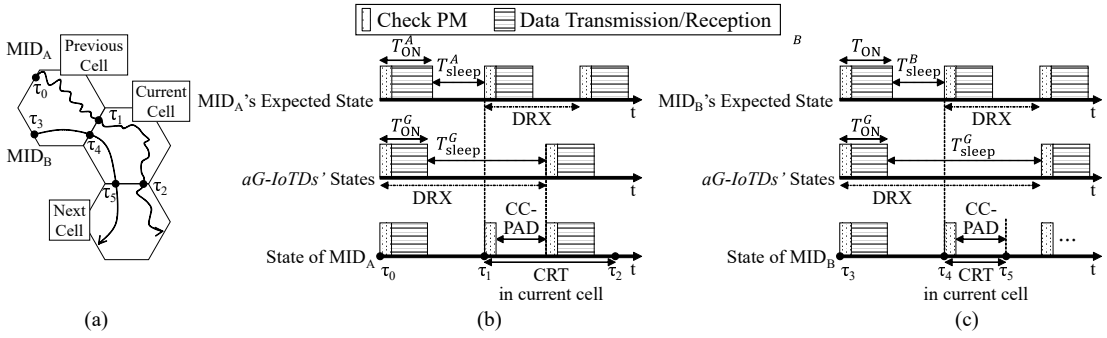

Figure 3.1: (a) Moving MIDs and their CRTs, (b) Case A: CC-PAD due to joining group during sleep time and CRT is greater than next G-DRX, and (c) Case B: CC-PAD due to joining group during sleep time and CRT is less than next G-DRX.

device (MID_A) starts walking at time instant τ_0 from the previous cell and enters into a new cell, i.e., its current cell at time instant τ_1, as shown in Fig. 3.1(a). Hence, MID_A resides within the previous cell for an interval of $\tau_1 - \tau_0$, which is defined as Cell Residence Time (CRT) in this paper. Similarly, in the current cell, MID_A has a CRT of $\tau_2 - \tau_1$. Besides, DRX of each IoT device has two states, namely, ON state (checks PM, RACH procedure, inactivity timer, and data transmission/reception) and sleep state. In Fig. 3.1(b), T^{A}_{ON} and $T^{\text{A}}_{\text{sleep}}$ are the expected ON time and sleep time of MID_A during its respective state before grouping. Similarly, in Fig. 3.1(c), T^{B}_{ON} and $T^{\text{B}}_{\text{sleep}}$ are the expected ON time and sleep time of MID_B. However, the *aG-IoTDs* have the same DRX based on group paging constraint. Therefore, we consider a common ON and sleep states of *aG-IoTDs* which is called as Group-ON time and Group-sleep time, and denoted as T^{G}_{ON} and T^{G}_{sleep} in this paper. T^{G}_{ON} and T^{G}_{sleep} are formulated as

$$T^{G}_{\text{ON}} = \max(T^{1}_{\text{ON}}, T^{2}_{\text{ON}}, \ldots, T^{N}_{\text{ON}}), \tag{3.1}$$

$$T^{G}_{\text{sleep}} = GDX_G - T^{G}_{\text{ON}}, \tag{3.2}$$

where T^{l}_{ON} is ON time of *aG-IoTD* l, and $l \in \{1, 2, 3, \ldots, N\}$ in the group G. T^{G}_{ON} is maximum period when at least one *aG-IoTD* is active and T^{G}_{sleep} is a period when all *aG-IoTDs* are in sleep mode. Because MIDs can receive the new group configuration through system information blocks broadcasted by eNB after wakeup in their expected state, MIDs can join a group either in Group-ON time or Group-sleep time. Moreover, DRX of a MID can change when it joins a group owing to group paging constraints. Hence, due to change in the DRX parameter, MIDs can receive PM if they join a group during Group-ON time and initiate data transmission/reception process. Otherwise, MIDs need to wait for the next wake

up of *aG-IoTDs* for PM as shown in Fig. 3.1(b) and Fig. 3.1(c). Thus, this waiting time can cause PAD for MIDs while joining a new group in the current cell. In this paper, we consider the PAD experienced by a MID in the current cell only because the configurations of groups in the next cell is unknown. Therefore, PAD is defined as the arrival delay experienced by a packet of a MID within its current cell due to change in DRX parameters after joining a group and is abbreviated as Current Cell PAD (CC-PAD) in this paper. It can be defined as the waiting time to receive the next PM in the current cell. We consider CC-PAD for uplink packets. However, it can be applicable for downlink packets. Moreover, CRT of a MID in the current cell can either be greater or less than the next wake-up of *aG-IoTDs* as illustrated in Fig. 3.1(b) and Fig. 3.1(c), respectively. For instance, in Fig. 3.1(b), CRT of MID_A $(\tau_2 - \tau_1)$ in the current cell is greater than the next wake-up whereas in Fig. 3.1(c), CRT of MID_B $(\tau_5 - \tau_4)$ in the current cell is lower than the next wake-up of *aG-IoTDs*. Here, MID_B can be an IoT device from a smart car which has high mobility and short CRT. In the former case, CC-PAD is the waiting time for the next PM of the group. However, in the latter case, MID misses the next PM in the current cell and look for next PM in the next cell, which is unknown. Hence, in this case, CC-PAD experienced while joining a group G in the current cell is equal to CRT of MID in the current cell. It has to be noted that Fig. 3.1(b) and Fig. 3.1(c) show the two cases that CRT ends during ON phase of MIDs because they do not perform cell reselection during sleep time, which is also considered in [13].

Moreover, each MIDs have their own PAD and PLR requirements based on application. MIDs can have packet loss if CC-PAD crosses the required PAD [14]. Thus, CC-PAD and PLR of MIDs should be satisfied while group paging. However, MIDs can avoid PLR challenge if PAD requirement is satisfied while grouping. Hence, PLR constraint can be relaxed in the proposed solution which

is presented in the further section. We model CC-PAD with PLR to show the relationship of CC-PAD, PLR, mobility and GDX_G in the next section. Based on the ongoing discussion, it can be observed that energy consumption depends on GDX_G and GF_G while QoS parameters depend on GDX_G. Therefore, we propose an approach to find the optimal GDX_G and GF_G such that energy consumption is minimized while meeting PAD and PLR requirements. The problem formulation and proposed models are presented in the subsequent sections.

3.3 Proposed Models

This section starts with the optimization problem formulation for grouping of MIDs with *aG-IoTDs*. Furthermore, we propose a novel mathematical model to analyze the CC-PAD and PLR experienced by MIDs while joining any new group in a cell. Thereafter, we propose an energy consumption model for group paging mechanism. Finally, we provide the solution of optimization problem and determine optimal group for MIDs.

3.3.1 Problem Formulation

Suppose the number of MID arrivals over a period of time at a cell follows a Poisson Process with average arrival rate of λ per unit time. Let M be the expected number of MIDs arrive at a cell within time t_a, where $M = \lambda t_a$. Regarding the reserved resources, the total available resources for the contention based RACH process are defined as Random Access Opportunities (RAOs) [8]. RAOs are equal to the number of the reserved frequency band in an RA slot multiplied by the number of available preambles for contention-based RACH. Without loss of generality, in this paper, we assume one frequency band in an RA slot. Hence, RAOs are equal to the number of preambles and has been denoted as R. How-

ever, the RACH process by more than R number of *aG-IoTDs* and MIDs leads to an increase in the probability of preamble collision. The collided devices perform random back-off and retransmit the preambles again with a ramp in power, which increases the energy consumption. Therefore, in this paper, we assume that a group can accept a maximum R number of IoT devices. This restriction is a group capacity constraint and has been mentioned as constraint C1 in (3.5). E is the total energy consumption of a group after grouping of M MIDs. Each MID m $(m = 1, \ldots, M)$ has PAD requirement d_m and an estimated CC-PAD based on the proposed delay model, represented as D_m. Similarly, the allowable PLR is Q_m and calculated PLR based on the proposed model is PL_m. Moreover, MIDs should meet the group paging constraints. Let MDX_m and MF_m be the required DRX and DTF for the m^{th} MID while GDX_G and GF_G be the G-DRX and G-DTF of joining group. The new G-DRX and G-DTF of *aG-IoTDs* and MIDs after grouping are denoted as $CGDX_G$ and CGF_G, respectively, and is formulated in (3.3) and (3.4), respectively.

$$CGDX_G = \min(GDX_G, MDX_1, \ldots, MDX_M), \tag{3.3}$$

$$CGF_G = \max(GF_G, MF_1, \ldots, MF_M). \tag{3.4}$$

The energy consumption of a group after grouping of MIDs depends on the $CGDX_G$ and CGF_G as explained in section 3.2.1. Moreover, the CC-PAD and PLR depends on the GDX_G as shown in section 3.2.2. Based on all the requirements and the objective of minimizing the energy consumption of IoT devices in

group paging, the optimization problem can be formulated as (3.5):

$$\begin{aligned} \underset{GDX_G, GF_G}{\text{Minimize}} \quad & E(CGDX_G, CGF_G), \\ \text{Subject to:} \quad & C1 : N + M \leq R, \\ & C2 : D_m \leq d_m, \\ & C3 : PL_m \leq Q_m, \\ & C4 : CGDX_G \leq MDX_m, \\ & C5 : CGF_G \geq MF_m. \end{aligned} \tag{3.5}$$

However, without loss of generality, we suppose that each group can have capacity to accept M MIDs, i.e., $R - N \geq M$, where N is the number of *aGIoTDs*. Therefore, we relax constraint C1 from (3.5). Constraint C2 and C3 represent constraints on PAD and PLR requirement of m^{th} MID, respectively, where $m \in \{1, 2, 3, ..., M\}$. D_m and PL_m depends on the GDX_G. Without loss of generality, we consider $d_1 \leq d_2 \leq d_3 \leq \cdots \leq d_M$ and $Q_1 \leq Q_2 \leq Q_3 \leq \cdots \leq Q_M$. Hence, constraint C2 and C3 can be changed to $D_m \leq d_1$ and $PL_m \leq Q_1$. This ensures that if MIDs meet the minimum PAD requirement (d_1), then they will meet their respective requirements also. Similarly, PLR requirement of MIDs can be set as a minimum PLR requirement (Q_1) of the MIDs. The CC-PAD experienced by M MIDs joining group G having GDX_G is same. Therefore, D_m can be replaced with D. Likewise, PLR for M MIDs is also the same, which is denoted as PL. Therefore, constraint C2 and C3 can be expressed in terms of d, D, Q and PL, where $d = d_1$ and $Q = Q_1$. We can rewrite the problem in (3.5)

as depicted in (3.6):

$$\begin{aligned} \underset{GDX_G, GF_G}{\text{Minimize}} \quad & E(CGDX_G, CGF_G), \\ \text{Subject to:} \quad & C1 : D \leq d, \\ & C2 : PL \leq Q, \\ & C3 : CGDX_G \leq MDX_m, \\ & C4 : CGF_G \geq MF_m. \end{aligned} \tag{3.6}$$

We can see from (3.3), (3.4), and (3.6) that energy consumption, PAD and PLR of MIDs depend on GDX_G and GF_G. Therefore, we need to find optimal GDX_G^* and GF_G^* for grouping of MIDs. In the next section, we propose a novel model to estimate the D and PL of (3.6).

3.3.2 CC-PAD and PLR Model

The CC-PAD depends on two factors, 1) whether MID joins a group during its sleep time or otherwise, 2) whether the CRT is less or greater than the next wake up of the group. GDX_G is represented as x in this model. In LTE, time is represented as frames and a frame is divided into subframes where duration of a frame and subframe is equal to 10 ms and 1ms respectively. Hence, in this paper, we represent all time variables in subframes and, therefore, time parameters are supposed to be discrete random variables. Moreover, CRT of a MID follows exponential distribution function with $1/\eta$ average cell crossing rate [27]. The Probability Density Function (PDF) of CRT having t_c subframes is expressed in (3.7):

$$P_{\text{CRT}}(t_c) = \eta e^{-\eta t_c}. \tag{3.7}$$

Here t_c is also discrete random variables and can be represented in subframes. We divide the CC-PAD and PLR model into two submodels: Model A and B. Model A and B estimate the PDF and Cumulative Distribution Function (CDF) of having CC-PAD owing to Case A and Case B respectively. Based on these two models, we calculate the total PDF of having CC-PAD. Thereafter, we estimate the CDF of CC-PAD that is greater than the allowable PAD, which provides PLR.

Model A: The time interval of DRX can be represented into radio subframes, and each subframe is equal to a millisecond. Moreover, in this model, t_c is the CRT of the previous cell. Suppose that a MID joins a group in a new cell at subframe t_j of a certain DRX cycle, which lies between 0 to $x-1$. The t_j can be equal to the remainder r when t_c is divided by x. Thus, the range of r is 0 to $x-1$, which gives the joining subframe. $r=0$ means t_c is multiple of x. $r=0$ represents a MID that joins the group at a subframe which is a multiple of x, such as x, $2x$, $3x$ and so on. Similarly, $r=1$ indicates that t_c is multiple of x with an addition of 1, such as $x+1$, $2x+1$ and so on. Likewise, $r=t_j$ means t_c is multiple of x with an addition of t_j, which provides the joining subframe at any G-DRX cycle. Hence, t_j depends on length of t_c and x. However, t_c is assumed to be a stochastic process and is defined as PDF in this paper from (3.7). Therefore, the r and t_j are also estimated as PDF. Let $P_{\text{join}}(t_j)$ be the PDF of a MID that joins t_j subframe and $P_{\text{rem}}(r)$ be the PDF that remainder is r. $P_{\text{join}}(t_j)$ should be equal to $P_{\text{rem}}(r)$, as shown in (3.8).

$$P_{\text{join}}(0 \leq t_j \leq x-1) = P_{\text{rem}}(0 \leq r \leq x-1). \tag{3.8}$$

The derivation of $P_{\text{rem}}(r)$ is expressed in (3.9).

$$\begin{aligned} P_{\text{rem}}(r=0) &= P_{\text{CRT}}(t_c = x) + P_{\text{CRT}}(t_c = 2x) + \cdots, \\ P_{\text{rem}}(r=1) &= P_{\text{CRT}}(t_c = x+1) + P_{\text{CRT}}(t_c = 2x+1) + \cdots, \\ P_{\text{rem}}(r=t_j) &= P_{\text{CRT}}(t_c = x+t_j) + P_{\text{CRT}}(t_c = 2x+t_j) + \cdots. \end{aligned} \tag{3.9}$$

Following first and second equation of (3.9), PDF of a MID that joins the group at t_j is estimated. Using (3.7), (3.8) and (3.9), the final expression of $P_{\text{join}}(t_j)$ can be written as

$$P_{\text{join}}(t_j) = P_{\text{rem}}(r=t_j) = \sum_{k=1}^{\infty} \eta e^{-\eta(kx+t_j)}. \tag{3.10}$$

However, $P_{\text{join}}(t_j)$ should be normalized such that area under the range of $1 \leq t_j \leq x-1$ is 1. let's A is the normalizing constant of $P_{\text{join}}(t_j)$. So, $P_{\text{join}}(t_j)$ expressed as

$$P_{\text{join}}(t_j) = \sum_{k=1}^{\infty} A * \eta e^{-\eta(kx+t_j)}. \tag{3.11}$$

The CDF of $P_{\text{join}}(1 \leq t_j \leq x-1)$ should be 1. We have

$$\sum_{t_j=1}^{x-1} \sum_{k=1}^{\infty} A * \eta e^{-\eta(kx+t_j)} = 1. \tag{3.12}$$

We can get A after solving (3.12). MIDs experience CC-PAD if they join during sleep time of the group and needs to wait till the next wakeup to receive PM for data transmission or reception as shown in Fig. 3 (b). The sleep time varies from $T_{\text{ON}}^G + 1$ (after ON time) to x, where T_{ON}^G is the Group-ON time, and derived from (3.1). Hence, the CDF of joining the group during sleep time of a

group is equal to CDF of $P_{\text{join}}(t_j)$

$$P_{\text{sleep}} = \sum_{t_j=T_{\text{ON}}^G+1}^{x-1} P_{\text{join}}(t_j), \tag{3.13}$$

where t_j ranges from $T_{\text{ON}}^G + 1$ to x. Let t_{A} be the CC-PAD casued when a MID joins the group at subframe t_j, which lies from $T_{\text{ON}}^G + 1$ to x, i.e., Group-sleep time. For instance, if a MID joins group at the first subframe of sleep time, i.e., $t_j = T_{\text{ON}}^G + 1$, then t_{A} is $x - (T_{\text{ON}}^G + 1)$. Thus, t_{A} is equal to $x - t_j$. t_{A} is function of of t_j. The calculation of t_j subframe is a stochastic process as shown in (3.10); therefore, the calculation of t_{A} is also stochastic. The PDF of having t_{A} is denoted by $P_{\text{delay}}^{\text{A}}(t_{\text{A}})$. However, t_{A} equals to 0 if $t_j \leq T$. So, $P_{\text{delay}}^{\text{A}}(t_{\text{A}}))$ is given by

$$P_{\text{delay}}^{\text{A}}(t_{\text{A}} = x - t_j) = P_{\text{join}}(t_j). \tag{3.14}$$

Substituting the value of $t_j = x - t_{\text{A}}$ and the value of P_{join} from (3.11) in (3.14) gives

$$P_{\text{delay}}^{\text{A}}(t_{\text{A}}) = \sum_{k=1}^{\infty} A * \eta e^{-\eta(kx+x-t_{\text{A}})}. \tag{3.15}$$

The CDF of $P_{\text{delay}}^{\text{A}}(t_{\text{A}})$ is defined below

$$CP_{\text{delay}}^{\text{A}}(t_{\text{A}}) = \begin{cases} \sum_{t_j=1}^{T} P_{\text{join}}(t_j) & t_{\text{A}} = 0 \\ \sum_{t_j=T+1}^{x-1} P_{\text{join}}(t_j) & 0 < t_{\text{A}} \leq x - (T_{\text{ON}}^G + 1) \end{cases}. \tag{3.16}$$

The CDF that a MID has t_{A} greater than d is expressed as the CDF of having CC-PAD from $d+1$ to $x - (T_{\text{ON}}^G + 1)$ and expression is shown below

$$CP_{\text{delay}}^{\text{A}}(t_{\text{A}} > d) = \sum_{t_{\text{A}}=d+1}^{x-(T_{\text{ON}}^G+1)} \sum_{k=1}^{\infty} A * \eta e^{-\eta(kx+x-t_{\text{A}})}. \tag{3.17}$$

Based on (3.16) and (3.17), we can see that CC-PAD depends on x.

Model B: A MID can miss next data transmission in new cell if CRT is less than the next DRX or wakeup. In this case, the CC-PAD experienced by a MID during joinging to a group in its current cell is the residence time in the cell, can be expressed as $t_{\mathrm{B}} = t_c$ and also shown in Fig. 3 (c), where t_c is the CRT in current cell and t_{B} is the CC-PAD owing to Case B. However, CRT is a stochastic process as expressed in (3.7). Therefore, we estimate CC-PAD in this case as stochastic process also. Let $P^{\mathrm{B}}_{\mathrm{delay}}(t_{\mathrm{B}})$ be the PDF of having CC-PAD owing to Case B. The CDF that a MID joins a group during sleep is P_{sleep} is expressed in (3.13). The PDF of having CC-PAD by a MID in Case B is equal to the product of CDF of having CRT less than or equal to $x - t_j$ and P_{sleep}. The expression is

$$P^{\mathrm{B}}_{\mathrm{delay}}(t_{\mathrm{B}} \leq x - t_j) = P_{\mathrm{sleep}} \times P_{\mathrm{CRT}}(t_c = t_{\mathrm{B}} \leq x - t_j). \tag{3.18}$$

Moreover, the CDF of having CRT less than $x - t_j$ is:

$$P_{\mathrm{CRT}}(t_c = t_{\mathrm{B}} \leq x - t_j) = \sum_{t_c=1}^{x-t_j} \eta e^{(-\eta t_{\mathrm{c}})}. \tag{3.19}$$

Substituting (3.13) and (3.19) in (3.18), we get

$$P^{\mathrm{B}}_{\mathrm{delay}}(t_{\mathrm{B}}) = \sum_{t_j=T^G_{\mathrm{ON}}+1}^{x-1} \sum_{k=1}^{\infty} A * \eta e^{-\eta(kx+t_j)} \times \sum_{t_B=1}^{x-t_j} \eta e^{(-\eta t_{\mathrm{B}})}. \tag{3.20}$$

The CDF that a MID has CC-PAD greater than the required PAD, i.e., $t_{\mathrm{B}} > d$ owing to this case is

$$P^{\mathrm{B}}_{\mathrm{delay}}(t_{\mathrm{B}} > d) = 1 - \sum_{t_B=1}^{d} P^{\mathrm{B}}_{\mathrm{delay}}(t_{\mathrm{B}}). \tag{3.21}$$

The CC-PAD when a MID joins a group occurs in two scenarios, as explained and modeled above. (3.15) and (3.20) are PDF of having CC-PAD due to Case A and Case B, respectively. The CDF of having CC-PAD greater than required delay d in the above two scenarios can derived from (3.17) and (3.21). That is

$$P_{\text{delay}}(t > d) = (P^{\text{A}}_{\text{delay}}(t > d) + P^{\text{B}}_{\text{delay}}(t > d))/2. \tag{3.22}$$

t_{A} in (3.17) and t_{B} in (3.21) are the CC-PAD variables in Case A and Case B, respectively, which is replaced by t in (3.22). The MID can have packet loss if the CC-PAD greater than the required PAD [14]. Hence, PLR is equal to the CDF of having PAD greater than the required PAD. Therefore, PLR exprienced by MID while joining a group in a new cell is equal to $P_{\text{delay}}(t > d)$ from (3.22). The PLR can be expressed as,

$$PLR = P_{\text{delay}}(t > d). \tag{3.23}$$

We have PDF of having CC-PAD in two cases A and B. The total PDF of having CC-PAD t when a MID joins a group can be expressed as

$$P_{\text{delay}}(t) = (P^{\text{A}}_{\text{delay}}(t) + P^{\text{B}}_{\text{delay}}(t))/2. \tag{3.24}$$

(3.23) and (3.24) are used to check the QoS requirement namely PAD and PLR while minimizing energy and their corresponding models are delineated in subsequent sections.

3.3.3 Energy Consumption Model

The energy consumption of MIDs and *aG-IoTDs* after grouping depends on DRX and DTF of MIDs and *aG-IoTDs* in a group. MIDs and *aG-IoTDs* of a group

after grouping are termed as *MG-IoTDs* in this paper. DRX allows *MG-IoTDs* to wake up and check for a PM. *MG-IoTDs* of a group consume E_{CPM} energy at each wakeup for checking PM. Hence, energy consumption of *MG-IoTD* i at wake-ups to check PM within time t_f can be expressed

$$E^i_{\text{CPM}} = E_{\text{CPM}} \times \frac{t_f}{CGDX_G}, \tag{3.25}$$

where t_f is the specified period of times to estimate DTF and number of wake-ups and $i \in \{1, 2, ..., N + M\}$ denotes *MG-IoTD*. The total energy consumption during a data transmission by i^{th} *MG-IoTD* depends on the ON time (random access time and data transmission) of *MG-IoTDs* (T^i_{ON}) after receiving the group PM. Let us consider that the energy consumption during ON time at each uplink data transmission is E_{freq}. E_{freq} is equal to the sum of energy consumption at RACH process (E_{RACH}) and power consumption of data transmission at subframe k (P^k_{DT}) times the total transmission time (T_{DT}).

$$E_{\text{freq}} = E_{\text{RACH}} + P^k_{DT} \times T_{DT}. \tag{3.26}$$

$$T_{DT} = P_i/C \times \mu. \tag{3.27}$$

Here, P_i is the packet size of data transmitted by i^{th} *MG-IoTD* of the group, C is the channel capacity and μ is the number of resource blocks assigned to a *MG-IoTD*. We consider that each *MG-IoTD* is allocated one resource block for data transmission in a subframe. The channel capacity can be estimated by Shannon-Hartley theorem.

Energy Consumption at RACH process: The energy consumption in RACH process involves energy consumption in all the four steps of RACH process. P_{PRACH} is a power consumption of signal from a *MG-IoTD* of group to eNB such

as preamble transmission at first step. That is

$$P_{PRACH} = \min(P_{CMAX}, PRTP + PL). \tag{3.28}$$

P_{CMAX} is the maximum transmission power of a *MG-IoTD* of a group. PL is the path loss represented in dB. $PRTP$ is the preamble received target power, which is the received power level of preamble at eNB. The $PRTP$ is expressed as

$$PRTP = PIRTP + \Delta_{prmbl} + (n_{tr} - 1) \times PRS. \tag{3.29}$$

$PIRTP$ is the power initial received target power, that is the initial value at which preamble is transmitted at start. Δ_{prmbl} is the preamble offset and its value depends upon the preamble format, for instance, 0 dB for format 0. n_{tr} and PRS are the current number of transmission of preamble by a *MG-IoTD* and power ramping step, respectively. The energy consumption in the second step has waiting time of RAR message is $T_{\text{RAR}}P_{\text{wait}}$ where T_{RAR} is the waiting time when eNB processes the preamble in subframe unit and P_{wait} is the power consumption when device is waiting for a RA slot/subframe. The energy consumption due to a IoT device receiving the Msg2 is $P_{\text{rx}}T_{\text{msg2}}$, where P_{rx} is the power consumption of receiving message from eNB in a RA slot. The energy consumption for Msg3 and Msg4 are $P_{PRACH} \times 1 + (T_{HARQ} + 1)P_{\text{rx}}$ and P_{rx}, respectively, where T_{HARQ} is the waiting time for acknowledgment from eNB. However, we have considered the maximum number of *MG-IoTDs* in a group to be equal to the number of available preambles R, which avoids the premable collison. Therefore, we are not including the power consumption at each retransmission of preambles. Accordingly, the

energy consumption of *MG-IoTDs* in a RACH process is:

$$E_{RACH} = T_{\text{RAR}} P_{\text{wait}} + P_{\text{rx}} T_{\text{msg2}} + P_{PRACH} + (T_{HARQ} + 1) P_{\text{rx}} + P_{\text{rx}}. \tag{3.30}$$

Power Consumption during Data Transmission at Subframe k:

$$P_{DT}^{k} = \min(P_{CMAX}, 10 \log_{10}(M_{PUSCH}(k)) + P_{PUSCH(j)} + \alpha(j) PL). \tag{3.31}$$

P_{DT}^{k} is the power consumption at subframe k. $10 \log_{10}(M_{PUSCH}(k))$ is the number of resource blocks at subframe k and $P_{PUSCH(j)}$ is the target eNB reception power. $\alpha(j)PL$ is path loss during data transmission. The energy consumption by a *MG-IoTD* i during uplink data transmission within t_f is

$$E_{\text{DT}}^{i} = E_{\text{freq}} \times CGF_{G}. \tag{3.32}$$

The total energy consumption of a group due to DRX and DTF is

$$E = \sum_{i=1}^{N+M} (E_{\text{CPM}}^{i} + E_{\text{DT}}^{i}). \tag{3.33}$$

However, the grouping of MIDs to a new group upon change of a cell can increase the energy consumption of other devices in a group and itself, as shown in section 3.2.1. Therefore, minimizing the total energy consumption of a group within time t_f is our objective, which is expressed in (3.33).

3.3.4 Optimization Solution

In this section, we provide the solution of the optimization problem formulated in (3.6) based on the designed models in previous sections. The constraint C1 in (3.6) represents PAD requirement of newly arrived MIDs. However, we have calculated the probability of having CC-PAD exprienced by MIDs in the previous subsection. Therefore, we need to change the constraint C1 in terms of expected CC-PAD of a MID joining a group having G-DRX x. The expected CC-PAD is

$$\mathrm{E_D}(x) = \sum_{t=1}^{x-(T^G_{\mathrm{ON}}+1)} t \times P_{\mathrm{delay}}(t). \tag{3.34}$$

$P_{\mathrm{delay}}(t)$ can be derived from (3.24) and CC-PAD varies from starting of sleep time, i.e., $(T^G_{\mathrm{ON}}+1)$ to x. Hence, D in the constraint C1 of (3.6) can be replaced with $\mathrm{E_D}(x)$. From (3.3) and (3.4), the term $\min(GDX_G, MDX_1, \ldots, MDX_M)$ and $\max(GF_G, MF_1, \ldots, MF_M)$ is the $CGDX_G$ and CGF_G in constraints C3 and C4 of (3.6), respectively. Without loss generality, we consider $MDX_{\min}$ is the smallest DRX, and $MF_{\max}$ is the highest DTF among incoming M MIDs . MDX_m and MF_m in constraints C3 and C4 of (3.6) are required characteristics of the m^{th} MID, where m varies from 1 to M. If MIDs after grouping have DRX less than and equal to $MDX_{\min}$ then all incoming MIDs can meet the group paging constraint. Therefore, MDX_m can be replaced by $MDX_{\min}$ in C3 of (3.6). Similarly, if DTF of incoming MIDs have a DTF greater than or equal to $MF_{\max}$, all the incoming MIDs can meet the group paging constraint. Thus, the MF_m can be changed to $MF_{\max}$ in C4 of (3.6). Moreover, in this paper, we constrained MIDs to join a group having G-DRX less than $MDX_{\min}$ and G-DTF greater than $MF_{\max}$. This constraint on MIDs restricts *aG-IoTDs* to change their configurations, which save their energy and avoid degradation of their QoS. Hence, $CGDX_G$ is equal to GDX_G and CGF_G is equal to GF_G. GF_G is written

as y from here. The constraint C2 of (3.6) is PLR constraint, which can be satisfied if MIDs can meet the PAD requirement, i.e., constraint C1. Therefore, we can remove the constraint C2. Hence, replacing the changes in (3.6), the representation of the optimization problem is

$$\begin{aligned} \underset{x,y}{\text{Minimize}} \quad & F(x,y) = E(x,y), \\ \text{Subject to:} \quad & C1 : g_1(x,y) = d - \mathrm{E_D}(x) \geq 0, \\ & C2 : g_2(x,y) = MDX_{\min} - x \geq 0, \\ & C3 : g_3(x,y) = y - MF_{\max} \geq 0, \\ & x > 0, y > 0. \end{aligned} \tag{3.35}$$

We have written the above optimization problem in terms $F(x,y)$ as objective function subject to constraints g_1,g_2, and g_3 greater than or equal to 0. The domain of x and y is real numbers greater than 0. We need to find the optimal value of x and y such that MIDs have minimum energy consumption and yet to meet the QoS requirements. The optimal value of x and y are the optimal GDX^*_G and GF^*_G.

The objective function and constraints in (4.42) are non-linear functions because of having reciprocal and exponential functions, respectively. Moreover, the problem is a convex optimization which can be proved based on derivative test easily. Thus, the problem can be solved by Lagrangian approach and applying Karush-Kuhn-Tucker (KKT) conditions provides the global optimal solution for such non-linear and convex optimization equation. The Lagrangian function

$L(x, y)$ of the above optimization problem is

$$L(x, y) = F(x, y) + \sum_{j=1}^{3} \mu_j \times g_j, \ \forall x > 0, y > 0, \tag{3.36}$$

where μ_j are Lagrangian multipliers. The KKT conditions lead to the following conditions:

$$\frac{\partial L}{\partial x} = \frac{\partial F(x^*, y^*)}{\partial x} + \sum_{j=1}^{3} \mu_j \times (\frac{\partial g_j(x^*, y^*)}{\partial x}) = 0,$$

$$\frac{\partial L}{\partial y} = \frac{\partial F(x^*, y^*)}{\partial y} + \sum_{j=1}^{3} \mu_j \times (\frac{\partial g_j(x^*, y^*)}{\partial y}) = 0,$$

$$\mu_j \times g_j(x^*, y^*) = 0 \text{ for } j = 1, 2, 3, \tag{3.37}$$

$$\mu_j \leq 0 \text{ for } j = 1, 2, 3,$$

$$g_j(x^*, y^*) \geq 0 \text{ for } j = 1, 2, 3,$$

$$x^* > 0, y^* > 0.$$

Solving these conditions in (3.37) gives the optimal value of x^* and y^* which yield minimum energy consumption of IoT devices in group paging mechanism. Based on the obtained solution after solving the KKT conditions, we present the results which promise reduction in the energy consumption significantly while meeting QoS requirement in the next section.

3.4 Performance Analysis

In this section, we evaluate the performance of the proposed grouping approach of MIDs. We first define the parameter settings. Thereafter, we analyze the relationship between the total energy consumption of a group if M MIDs join a

group having various G-DRX and G-DTF in Section 3.4.2. Furthermore, based on the proposed CC-PAD model, we examine probability functions of having CC-PAD, expected CC-PAD and PLR experienced by a MID at different G-DRX for two types of mobility patterns, namely long-CRT and short-CRT. Finally, we compare the total energy consumption of group after joining of MIDs in group paging mechanism by our proposed approach with a random grouping approach.

3.4.1 Parameter Settings

We have considered that values of M ranges from 1 to 3 within $t_a = 1$ second at a cell having 2 km radius. Without loss of generality, we consider each group having $N = 10$ *aGIoTDs*, which should be less than 54 owing to limited number of preambles. The characteristics of an IoT device varies with IoT applications. The DTF can perform single data transmission over $t_f = 1$ hour for smart meter applications to that requiring 360 transmissions in an hour for fleet management applications [28]. DRX values of an IoT device can vary up to 84000 seconds which is the latest value as per extended DRX in LTE-A Pro networks [13]. However, in our analysis, DRX values of a MID have been varied between 10 seconds to 100 seconds and DTF in the range of 10 to 100. Furthermore, the average CRT of MID is set as 800 seconds and 144 seconds resembling motion of IoT devices during walking (with wearable devices) and driving cars in an urban area, respectively. Additionally, we analyze the group having G-DRX values varying between 10 to 200 seconds and G-DTF between 10 to 200. In our analysis, we have assumed that each group has capacity to accept incoming MIDs. The basic unit of time for channel resource allocation is in milliseconds (1 sub-frame). The available bandwidth for a group of devices is 20 MHz. There are around 200 resource blocks in the 20 MHz channel for data transmission, and each device has an allocation of a resource block at each subframe [29]. We adopt timing parameters associated

with the RACH from [16]. The packet size of IoT applications is usually small and can vary from 8 bytes to 100 kilobytes [30]. Furthermore, the parameter related to power consumption during RACH and data transmission process is also defined in [8]. For simplicity, we have considered constant path loss for all IoT devices. Moreover, the power consumption for checking of a PM is 250 mW/TTI, where TTI is Transmission Time Interval. The DRX of an IoT device is usually derived based on the PAD requirements. Hence, a device with long DRX value can be delay tolerable and have long PAD requirement [17, 18]. Therefore, we have considered delay tolerable MIDs such as waste management collection, smart logistics etc. with PAD requirement between 1 to 100 seconds [31, 32]. Some applications such as mobile environment monitoring can have delay tolerance of 30 minutes. Based on the above parameter settings, we carry out the mathematical analysis of CC-PAD and PLR model with various G-DRX and the proposed scheme.

3.4.2 Energy Consumption versus Characteristics of a Group

In Fig. 3.2, we vary the GDX_G while keeping GF_G fixed at 20 in order to observe the impact of G-DRX on the energy consumption of the group after joining of incoming MIDs. It can be observed from Fig. 3.2 that total energy consumption decreases till the minimum DRX value (indicated with arrows) of incoming M MIDs (M values from 1 to 3) because of the group paging constraints in (3.3) which change the MID's DRX to G-DRX. Hence, MIDs need to wake up more than required DRX in an hour which leads to an increase in total energy consumption. However, total energy consumption is constant with GDX_G greater than minimum DRX of MIDs due to group paging constraints. Similarly, in Fig. 3.3, we vary the GF_G while keeping GDX_G fixed at 10 seconds and M ranges from 1 to 3. As can be seen in Fig. 3.3, the total energy consumption is constant for

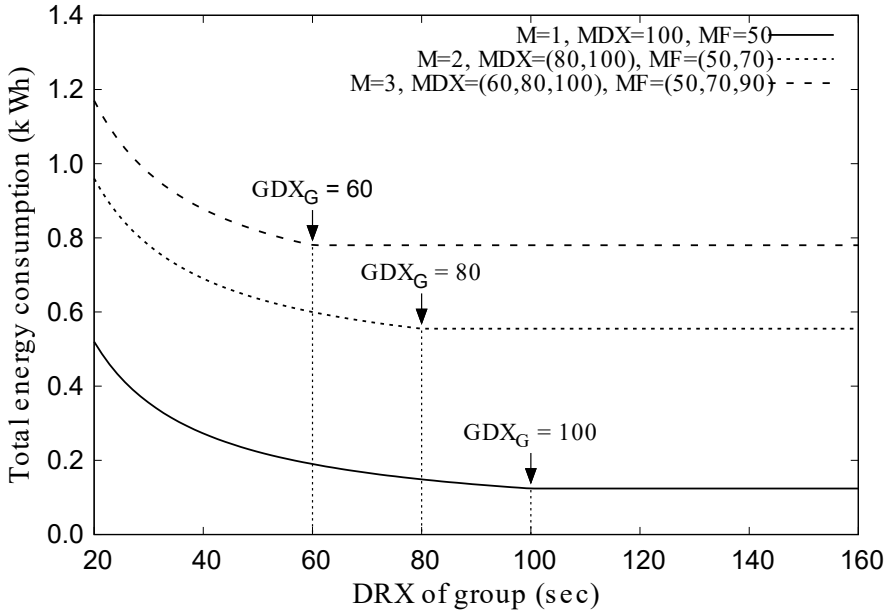

Figure 3.2: Total energy consumption of group after grouping MIDs at various GDX_G and $GF_G = 20$.

G-DTF having value less than or equal to the maximum DTF of incoming MIDs but increases significantly with a group having DTF more than the maximum DTF of MIDs (also indicated with arrows). The reason for such a behavior can be ascribed to the change in MIDs' DTF to the G-DTF when G-DTF is higher than their DTF to meet group paging constraints, as mentioned in (3.4). Hence, in this case, MIDs should join a group having optimal G-DTF less than and equal to the maximum DTF of incoming MIDs. Fig. 3.2 and Fig. 3.3 show evaluation of energy consumption at constant GF_G and GDX_G, respectively. We evaluate the total energy consumption of a group at various G-DRX and G-DTF configurations when $M = 3$, $MDX = (60, 80, 100)$ and $MF = (50, 70, 90)$ in Fig. 3.4. From Fig. 3.4, we can see that the energy consumption converges to a plane where energy consumption is minimum. That plane is shown within arrows in Fig. 3.4. The groups having characteristics within the plane can give minimum energy

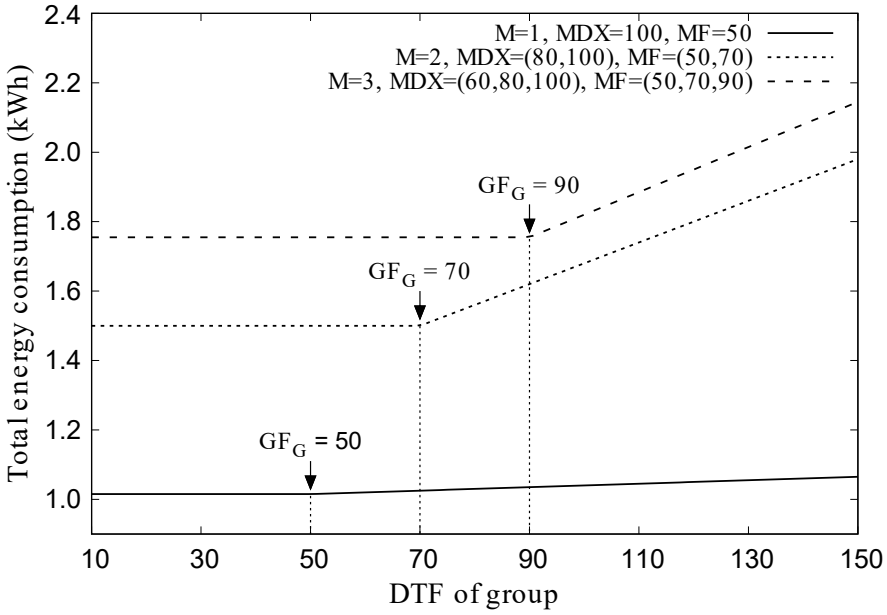

Figure 3.3: Total energy consumption of group after grouping MIDs at various GF_G and $GDX_G = 10$.

consumption. Hence, MIDs should join a group having characteristics within the plane.

However, the CC-PAD depends on the GDX_G as discussed in section 3.3.2. Hence, we need to assure that the MIDs should meet their PAD requirements at selected GDX_G. Fig. 3.2, Fig. 3.3, and Fig. 3.4 do not include the PAD requirements. Hence, it may be optimal if it meets the PAD requirements, otherwise, we need to choose the GDX_G which satisfies the PAD constraints. We show the optimal results in the subsequent section after studying the relationship between CC-PAD, PLR and GDX_G.

3.4.3 Relation of G-DRX with CC-PAD and PLR

Based on (3.16) and (3.20), we evaluate the CDF of having CC-PAD of a MID due to Case A and B under different G-DRX as depicted in Fig. 3.5 and Fig. 3.6,

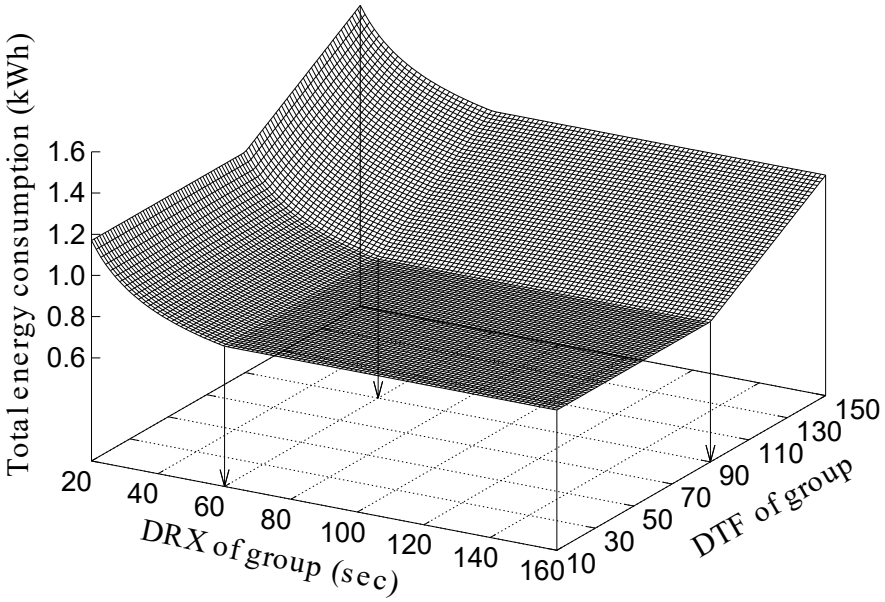

Figure 3.4: Total energy consumption of group after grouping MIDs at various GDX_G and GF_G.

respectively. It can be observed from Fig. 3.5 that CDF of having CC-PAD due to Case A increases with growing G-DRX. This is because a group with longer DRX has longer sleep time. Longer sleep time entails more chances of MIDs to join a group during sleep time. Furthermore, it can be discerned from Fig. 3.6 that CDF of having CC-PAD due to Case B increases with increase in G-DRX. The reason for such behaviour is the CDF of having CRT less than GDX_G as mentioned in (3.19) and is shown in Fig. 3.7. In Fig. 3.7, the CDF of having CRT less than GDX_G is low at short G-DRX because average CRT of MIDs is usually higher than short G-DRX such as 10 to 60 seconds. Hence, the chance of having average CRT less than short G-DRX is low. However, long G-DRX such as 200 second has more chance of having average CRT less than G-DRX because high-speed devices such as train or driverless car usually have average CRT around that value. Accordingly, we have shown the total probability of having CC-PAD

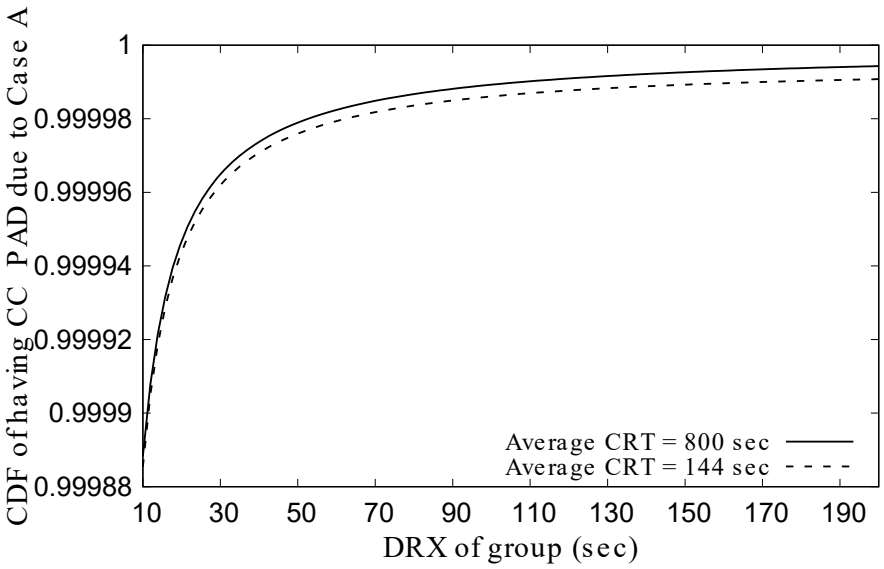

Figure 3.5: The CDF of having CC-PAD under different GDX_G due to Case A as mentioned in section (3.3.2).

because of Case A and B in Fig. 3.8. The total CDF of CC-PAD increases as G-DRX progress because the CDF of CC-PAD due to Case A and Case B also increase with the rise of G-DRX, as explained before and shown in Fig. 3.5 and Fig. 3.6. Moreover, we can also observe in Fig. 3.5 that long-CRT MIDs have a higher probability of having CC-PAD than short-CRT MIDs because former has high CDF of having CRT greater than G-DRX as can be seen in Fig. 3.9. Similarly, in Case B, short-CRT MIDs have a higher CDF of having CC-PAD than former because short-CRT MIDs have a high CDF of having CRT less than DRX of a group, as can be observed in Fig. 3.7. We can see similar nature for the combined probability at different mobility patterns. From the above discussion, it can be confirmed that CDF of having CC-PAD depends on G-DRX and mobility pattern of MIDs. To understand optimal behavior of CC-PAD from Fig. 3.5, Fig. 3.6 and Fig. 3.8, we estimate the expected CC-PAD when a MID joins a group having different GDX_G from (3.34) and examine it in Fig. 3.10. From

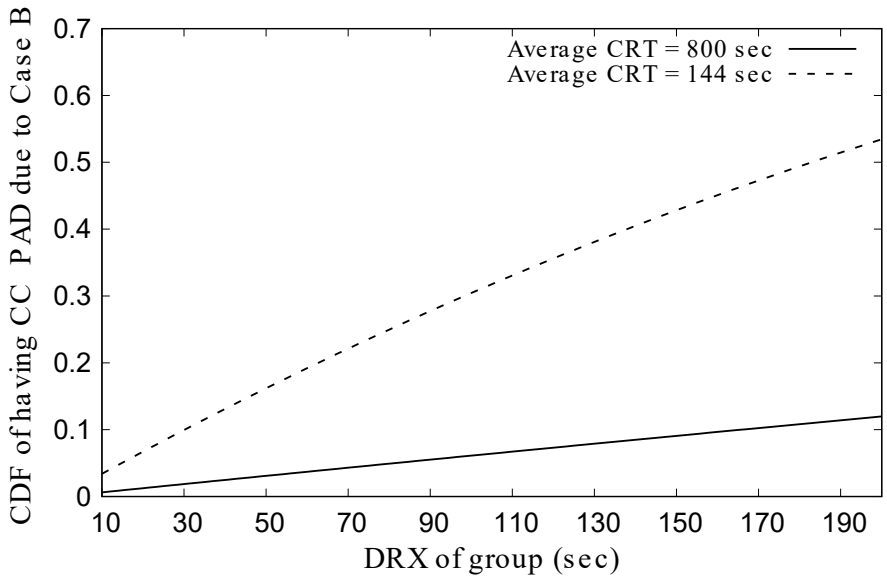

Figure 3.6: The CDF of having CC-PAD under different GDX_G due to Case B as mentioned in section (3.3.2).

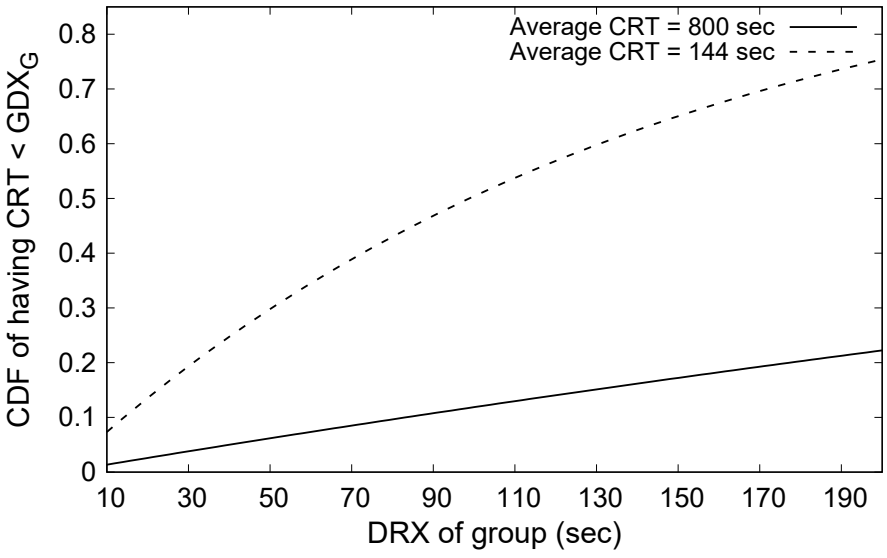

Figure 3.7: The CDF of CRT is less than GDX_G when $1/\eta = 144$ and 800 second.

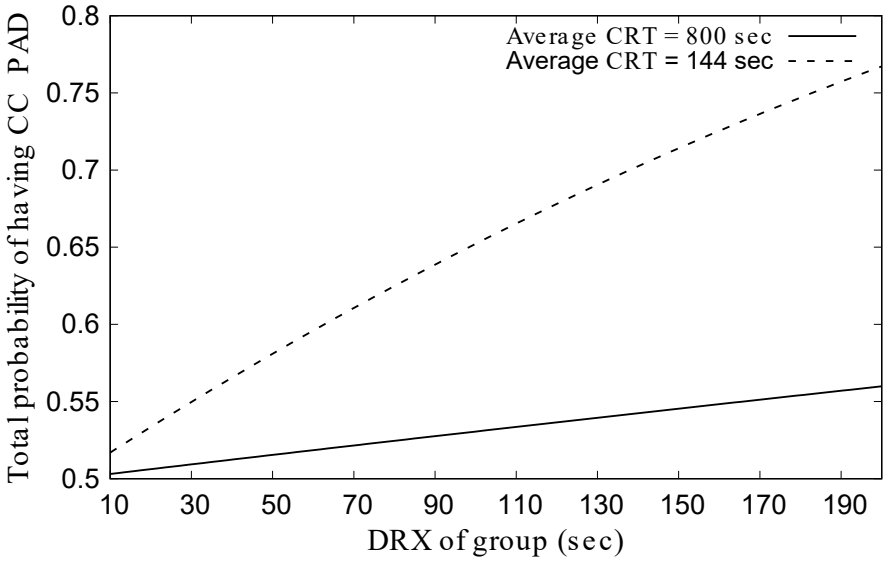

Figure 3.8: The total probability of having CC-PAD under different GDX_G.

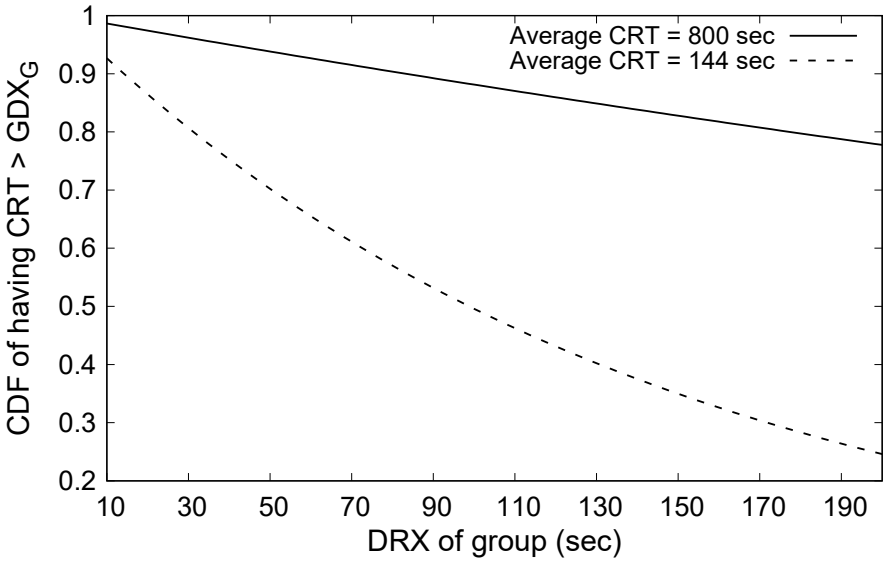

Figure 3.9: The CDF of CRT is greater than GDX_G when $1/\eta = 144$ and 800 second.

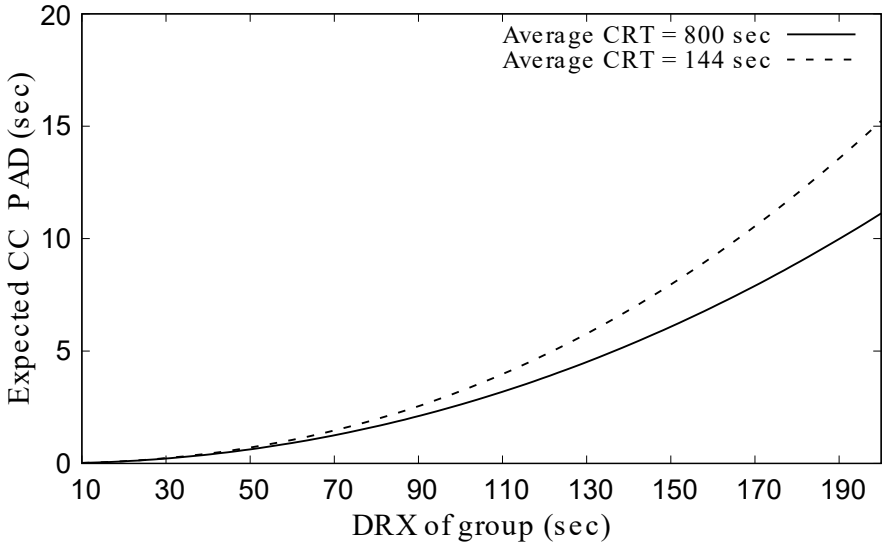

Figure 3.10: Expected CC-PAD at different GDX_G.

Fig. 3.10, we can observe that expected CC-PAD grows when GDX_G increases. The reason for the rise of expected CC-PAD is longer sleep time within a longer DRX and thus, high probability of joining the group during sleep time as discussed before. Moreover, the CC-PAD of MID having average CRT = 144 second has more expected CC-PAD than MID having average CRT = 800 second because faster MIDs have more probability to leave the cell before next wake up. Based on this evaluation, we choose GDX_G such that it can ensure that MIDs satisfy constraint C1 in (4.42). Finally, we study the PLR for two types of MIDs at various GDX_G in Fig. 3.11 having CC-PAD requirement of 10 seconds. Fig. 3.11 indicates that PLR grows with an increase in GDX_G. This happens because a MID experiences very less CC-PAD at low DRX implying that they will meet the CC-PAD requirements. Hence, PLR of MID also depends on GDX_G.

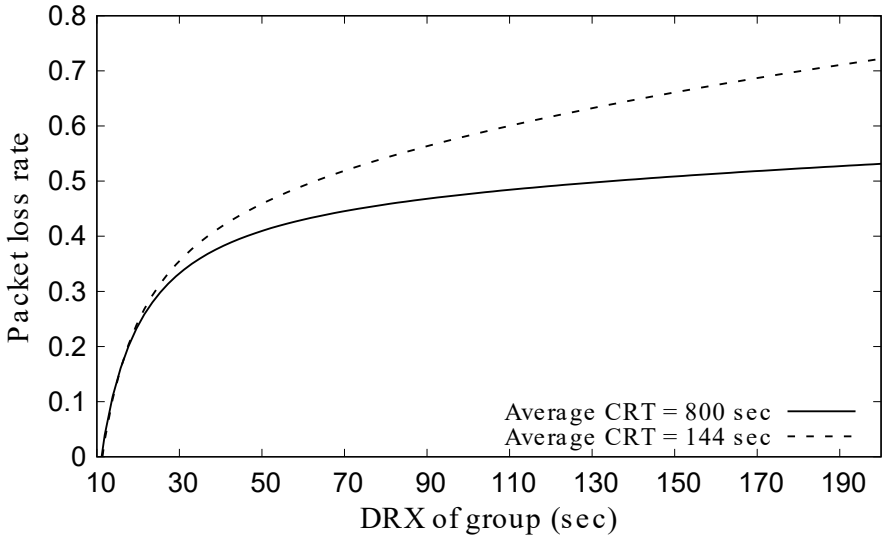

Figure 3.11: PLR at different GDX_G.

3.4.4 Proposed Approach versus Random Grouping Approach

Fig. 3.12 demonstrates the total energy consumption of a group using both random grouping and our proposed grouping method. Random grouping can be defined as an approach to group irrespective of IoT characteristics, for instance, based on location. Random grouping method is usually developed for traditional devices such as smartphones and laptops which do not possess specific characteristics like periodic data transmission since there are many applications executing in a single system, each of which can have different requirements. In random grouping approach, the MIDs join any group having DRX from 10 seconds to 200 seconds and DTF from 10 to 200. We repeat the process for 1000 times then average the total energy consumption in 1000 processes which gives total energy consumption in random grouping approach. In our proposed approach, we estimate optimal group having optimal GDX_G and GF_G based on solving the

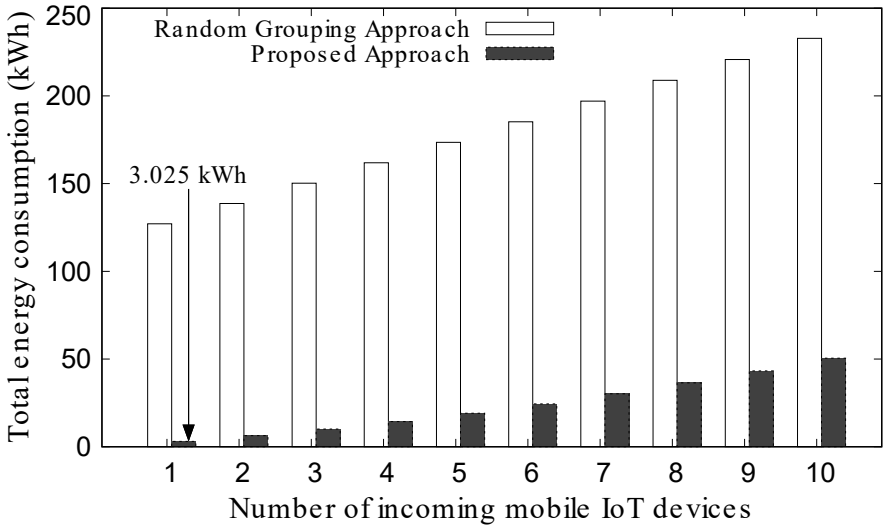

Figure 3.12: The total energy consumption of *MG-IoTDs* in a group using random grouping and proposed method.

KKT-conditions. From Fig. 3.12, we can observe that our proposed approach significantly reduces the hourly energy consumption with respect to random grouping. In addition, our proposed approach considers the QoS parameters such as PAD and PLR, which guarantees QoS.

3.5 Summary

In this chapter, we proposed a novel energy efficient grouping method for MIDs while considering their QoS requirements such as PAD and PLR. The proposed approach is based on two characteristics, namely, DRX and DTF of an existing group in a cell which is responsible for energy consumption of IoT devices. Moreover, the CC-PAD and PLR exprienced by MIDs depend on the DRX of a group. In this chapter, we have designed novel mathematical models to study the behavior of energy consumption, CC-PAD and PLR at various G-DRX and G-DTF of a group. Based on the mathematical analysis of energy consumption,

CC-PAD and PLR models, we have concluded that CC-PAD and PLR is low at low G-DRX of a group whereas energy consumption is high at low G-DRX and high G-DTF. The proposed scheme was found to outperform its random grouping counterpart. Additionally, based on our mathematical analysis, we have studied the QoS parameters for two types of mobility patterns at various G-DRX of a group which concluded that devices with short-CRT (high speed devices) have higher CC-PAD and PLR than their long-CRT (low speed devices) counterparts. We have studied limited types of mobility patterns in this paper but MIDs can have diverse patterns.

Chapter 4

A Novel Network-Aware Internet-wide Port Scan (IWPS)

4.1 Introduction

This chapter presents overview of considered security technology that is Internet Security Protocol (IPSec) to provide the information security such as confidentiality, integrity and so on. Furthermore, this chapter explains the research challenges of IWPS that may degrade the IoT security, rather than improving it. To address this challenge, this chapter proposes novel mathematical IEEE 802.11ah's network-aware IWPS models to evaluate risk on IoT devices for security maximization. Finally, numerical analysis of proposed models are presented in this chapter.

4.2 Considered Security Technology

The cyber security comprises of techniques to ensure protection of data CONFidentiality (CONF), INTegrity (INT), and AVAilability (AVA) with efficient

design of vulnerability scanning and developing intrusion protection programs. In this thesis, we consider IWPS for vulnerability scanning, which can provide intrusion detection programs to carry out necessary steps to fix the vulnerability. However, CONF, INT, and AVA are provided through encryption, authentication and avoiding distributed denial attacks (DDOS), respectively. All IoT devices will soon be IPv6 enabled, owing to the massive number of devices and the advantages of low-energy consumption [33, 34]. The Internet-security (IPsec) protocol is a promising technology for enabling end-to-end security of devices having low resources, owing to their integration into the Internet Protocol (IP) [35, 36]. Additionally, IPsec support is mandatory for IPv6-enabled IoT devices. Thus, we assume that most such devices will use the IPSec protocol as a standard solution.

There are two primary types of IPSec protocols: encapsulating security payload (ESP) and authentication header (AH). However, ESP includes both encryption and authentication, which provides both CONF and INT services. AH checks only INT of packets. Hence, we consider ESP in this thesis. IPSec also provides the flexibility to choose an encryption algorithm between two IPsec devices by negotiation. In IPsec, first, sec-admins define a set of transform sets that is a collection of encryption algorithms for devices. During the negotiation process, IoT devices check their matching transform sets, and matched sets are communicated through the security association. Thus, each pair of devices choose the encryption algorithm for ESP protocols based on chosen transform set. Each transform set define encryption algorithm, and are sorted according to the priority of use of encryption algorithm in a IoT device. However, security admin can define an encryption algorithm for each transform set based on the observation of traffic and quality of service (QoS) [37–39] requirement. For example, in [39], a flexible game-theory approach was proposed to determine multi-level security based on the trade-off between the encryption algorithms, the key length, and the network

throughput as a QoS parameter. Here, we consider IoT throughput as a QoS metric to set the encryption algorithm. Moreover, in this thesis, we consider different encryption-algorithm choices, including hardware (HW) implementations of data encryption standards (DES-HW), Triple DES (3DES), AES-128 bit, etc. However, heavyweight protocols such as AES-256 bit, cannot be implemented in IoT owing to its computational complexity. Hence, in this study, we assume only those algorithms that are applicable to IoT. Security of IoT is evaluated using risk metric which considers both vulnerability attack and information security services such CONF, INT, and AVAilability (AVA). Risk is defined as a security metric which determines level of risks on any device that is estimated from impact of vulnerability, threat, and asset-importance on security of a device. The risk on a device is formulated as follows [40].

$$Risk = V \times Th \times AI, \tag{4.1}$$

where V is the common vulnerability scoring-system (CVSS) score of the a vulnerability in a device. Th and AI are the threat level and asset weight of device, respectively. The threat can be defined in terms of access level to any device by an attacker such as physical access, remote access, and so on. The threat parameter is quantified as a weight value equals to the threat of actual access by an attacker compared to a high threat i.e. physical access. Hence, the threat value lies from 0 to 1. A threat value zero means there is no threat while threat value of one denotes a high threat due to easy physical access. As we know IoT devices are publicly available and can have easy physical access and also have a weak password. Hence, threat value is set to 1 in this thesis and which is also considered in [41]. Moreover, asset is defined as the importance of a device, which can cause huge loss to users in case of security breaches or impacts of a compromised device. It is quantified as weight value of the importance of a device compared to

a highly important device. Hence, the asset value of any device also varies from 0 to 1. IoT devices have high importance because any compromised IoT device can act as a bot agent and can initiate attacks on other network infrastructures, which can cause huge losses. Thus, asset value is set as 1 in this thesis, which is considered same in [41]. However, CVSS value varies from 0 to 10 [41, 42]. The CVSS metric defines severity levels for certain range of CVSS value. The qualitative severity scale of CVSS value are given as: CVSS=0 signifies no risk, 0.1 to 3.9 defines low risk, 4.0 to 6.9 means medium risk, 7.0 to 8.0 considers high risk and 9.0 to 10.0 represents critical risk. Thus, risk values can be quantified from 0 to 10 according according to defined CVSS ranges, and constant value of threat, and asset. The risk value such as 1 represents lower risk and high security whereas higher risk value such as 10 represents highly insecure devices. However, the CVSS is a dynamic scoring standard for any vulnerability and covers three metrics: base, temporal, and environmental scores [42]. The base score is the primary value of a vulnerability that is static with respect to time and the environment. Meanwhile, the temporal score reflects the variation in the base score over time owing to the attack likelihood, available patching solution, and report confidence. Moreover, the environmental score represents the variation of the base and temporal scores, owing to the fluctuation in the device's security environment, such as the variations and requirements of CONF, INT, and AVA. Therefore, in the following sections, we show research challenges concerning the impact of inappropriate scan rate on temporal and environmental scores, owing to the poor network performance of IEEE 802.11ah that can increase the risk on IoT devices.

4.3 Research Challenges: Security Degradation

This section presents the research challenges concerning the degradation of IoT security by impacting temporal and environmental score owing to network-oblivious IWPS.

4.3.0.1 Impact of Scan Rate on Temporal Score

The temporal score consists of remediation level, report confidence, and exploit code maturity (ECM) [42]. The remediation level of a vulnerability signifies whether an official patching solution is available. Because our focus is to study the network performance of IWPS, we assume that patching algorithms are available. Moreover, the report confidence signifies the confidence in the presence of vulnerabilities and the reliability of technical details. We also suppose that confirmed and technically sound reports on vulnerabilities exist owing to the contributions by reputed standards, such as common vulnerabilities and exposures [43]. However, ECM is evaluated on the basis of the attack likelihood (AL) of the vulnerability [42]. AL is a relative term that depends on the patching delay, defined as the time required to identify and fix a port vulnerability. For example, if sec-admins can regularly identify and patch vulnerabilities before a compromise, then attackers will not have chance to attack any IoT device. Moreover, the patching delay is a function of both the scan delay for probing each port for vulnerabilities and the time taken for processing the patching algorithms. The execution time of patching algorithms is considered constant since it is independent of the scan rate. The patching delay can be affected significantly by an increase in the scan delay, leading to long patching delay and high AL. Therefore, the scan delay should be minimized to reduce AL and impact on the temporal score. Moreover, the scan rate influences the scan delay. Both low and high scan rates can increase the scan delay. For example, the probing of all ports (around 65535) in each

IoT device per AP with a low scan rate takes a longer time; whereas a high scan rate can increase the scan delay because of the numerous backoffs in obtaining channel resources or scan packets drops, owing to the channel congestion in each RAW slot. Hence, a low and high scan rate can increase AL by having long patching delay, which increases the temporal score and the risk to an IoT device. Therefore, an optimal scan rate is required to minimize this risk. Moreover, the scan rate can influence the environmental score of CVSS, which is discussed in the next section.

4.3.0.2 Impact of Scan Rate on Environmental Score

The environmental metric represents the impact on the security of devices, owing to modification of the devices' security environment, such as CONF, INT, AVA, and their requirements. The CVSS determines the CONF, INT, and AVA requirements based on the importance of the affected device to an organization (e.g., service availability). However, the requirements of these parameters for IoT are always high, owing to easy access to IoT devices. However, the security is low, owing to the constrained resources. The compromise of a single security services in IoT can lead to attacks on other IT assets of the organization (e.g., DDoS) [7]. Therefore, in this study, we consider providing high CONF, INT, and AVA. The full influence of the environment metric is estimated using the modified base score, which is altered because of the variation in the CONF and INT protocols. The value CONF and INT at any time depends on the strength of encryption algorithm. As discussed in the system model that security admin decides encryption algorithm based on available throughput for each IoT device. High available IoT throughput can allow to set strong encryption whereas low throughput allows weak encryption algorithm. A high scan rate leads to channel congestion at any WLAN and resource exhaustion for IoT packets. Hence,

a high scan rate can cause low IoT throughput which in turn might impose the implementation of weak encryption algorithms. However, the size of scan packets is small and frequency of incoming scan packets are low. Though the scanning of an enormous number of ports for large number of devices per IEEE 802.11ah enabled AP can take several days or weeks owing to limited bandwidth availability in IEEE 802.11ah. During this scan duration, if security methods are not adjusted according to the scan rate then an IoT device can experience degraded services for several days or weeks, which can have serious impact on IoT services that require high throughput or low delay. Hence, security methods such as encryption algorithm should change at various scan rates. Based on these reasons, a high scan rate impacts the CONF and INT security services. Moreover, owing to high scan rate, network congestion may be responsible for packet loss if the packets are unable to obtain channel resources within their maximum number of permitted re-transmissions. Hence, a high scan rate also reduces the AVA of IoT packets and services. Such a scenario is defined as a scan attack. Hence, AVA should be improved to avoid transformation of a positive scan into an attack and the consequent security degradation. Therefore, in the next section, we propose a novel security and network model to optimize the scan rate to minimize the risk to IoT.

4.4 Proposed Models for Security Maximization

In this section, we propose a novel mathematical model in support of network-aware IWPSs evaluating the risk of a device. We divide the proposed model into two sub-models: IoT/scan throughput and network-aware risk evaluation. Finally, we formulate the optimization problem on the basis of these two models to estimate the optimal value of the scan rate needed to minimize the risk value.

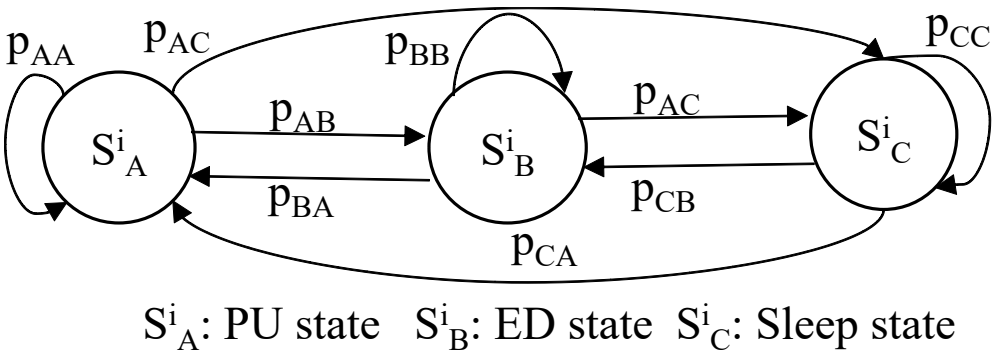

Figure 4.1: Markov model used to estimate traffic for IoT device i.

4.4.1 IoT and Scan Throughput Model

First, we present a novel queue model for port scanning and Het-IoT traffic. Then, we propose a mathematical model to estimate scan and IoT throughput over IEEE 802.11ah.

4.4.1.1 Queue Model

This chapter considers N_{AP} number of fully connected IEEE 802.11ah enabled access points (AP) over the Internet with an average of N number of IPsec and IPv6 enabled IoT devices per AP. In IEEE 802.11ah MAC protocol as shown in Fig. 2.4, a beacon interval (t_{b}) is divided into K RAW frames (RFrames), and length of each RFrame is $t_{\mathrm{R}} = t_{\mathrm{b}}/K$. These N devices access the wireless channel within a RAW frame. However, a RAW frame is slotted into RAW slots (Rslots) and length of each Rslot is t_{s}, as shown in Fig. 2.4. The RAW mechanism allows a group of devices to access the channel in a Rslot while allowing the other devices to sleep. Hence, there are M groups of devices with group size g. Each group accesses the channel using CSMA/CA process, as shown in Fig. 2.4. However, we consider the non-crossing case wherein an ongoing transmission is not allowed to cross the designated slot. To restrict transmission to another Rslot, IEEE

802.11ah introduces a holding time (t_{Ho}) at the end of the Rslot, which interrupts the ongoing transmission, as depicted in Fig. 2.4. Thus, the Rslot time available for transmission is t'_{s}=$t_{\mathrm{s}} - t_{\mathrm{Ho}}$.

There are two major types of traffic generated by IoT: periodic update (PU) and event-driven (ED) [44], which are uplink (UL)-dominant. To model such traffic from each IoT device, a semi-Markov model that includes PU, ED, and a payload exchange (PE) was proposed in [44]. PE packets include the data generated following PUs or EDs that provide additional details after an event. However, PE packets can be merged with PUs and EDs, because their arrivals coincide. Hence, we revise the traffic model to provide a more precise and realistic Het-IoT scenario. In Het-IoT traffic, each IoT device can have different packet-arrival rates. Therefore, in this model, we consider that the PU of IoT device i has $\lambda_i^{\mathrm{ULPU}}$ packet arrivals for transmission per unit time, where i varies from 1 to N, and N is the total number of devices per AP. In this study, the packet-arrival rate is defined per unit time, which is equal to a beacon interval. The periodic update downlink (DL)-packet arrival rate for device i in an AP is $\lambda_i^{\mathrm{DLPU}}$. We also assume that the arrival of the ED and the following DL control packets follow a Poisson process [44]. To estimate the total packet arrival for UL (including ED with periodic updates), we propose a Markov model, as shown in Fig. 4.1. Because IEEE-802.11ah-enabled IoT devices have a sleep mode, we consider the sleep state in our model. The Markov model in Fig. 4.1 represents the traffic model for device i. PU, ED, and Sleep are three states for device i, denoted as $\mathrm{S}_{\mathrm{A}}^i$, $\mathrm{S}_{\mathrm{B}}^i$, and $\mathrm{S}_{\mathrm{C}}^i$, respectively. Based on the Markov model, the average number of packet arrivals for UL (λ_i^{UL}) and DL (λ_i^{DL}) of device i per unit time is formulated

as follows:

$$\lambda_i^{\mathrm{UL}} = \lambda_i^{\mathrm{ULPU}} P_{\mathrm{A}} + \lambda_i^{\mathrm{ULB}} P_{\mathrm{B}}, \tag{4.2}$$

$$\lambda_i^{\mathrm{DL}} = \lambda_i^{\mathrm{DLPU}} P_{\mathrm{A}} + \lambda_i^{\mathrm{DLB}} P_{\mathrm{B}}, \tag{4.3}$$

where λ_i^{ULB} and λ_i^{DLB} are the respective packet-arrival rates for UL and DL (control packets) transmissions at $\mathrm{S}_{\mathrm{B}}^i$. However, there is zero packet generation during the sleep state. Hence, there is no arrival in the queue during sleep time. The DL IoT data packets are thus added to the AP queue. Let P_{A}, P_{B}, and P_{C} be the steady-state probabilities for states A, B, and C, respectively. These steady-state probabilities can be calculated by solving $\vec{\pi}T{=}\pi$, where $\vec{\pi}$ and T are the steady-state probability distribution vector and transition matrix, respectively, with constraint $P_{\mathrm{A}}{+}P_{\mathrm{B}}{+}P_{\mathrm{C}} = 1$. In this model, we consider the M/M/1:∞/first-in-first-out queue, which supports the periodic and Poisson processes. Let Y_i and Y_{AP} be the packet-processing time of device i and the AP, respectively. This is defined as the time required to successfully transmit a packet. Hence, based on the Markov queue model, the probability of having at least a packet in device i (Q_i^{Dnemp}) and the AP (Q^{APnemp}) queue during Y_i and Y_{AP} is given by

$$Q_i^{\mathrm{Dnemp}} = 1 - e^{-(\lambda_i^{\mathrm{UL}} + \lambda_i^{\mathrm{ACKscan}})Y_i}. \tag{4.4}$$

$$Q^{\mathrm{APnemp}} = 1 - e^{-(\sum_{i=1}^{N}(\lambda_i^{\mathrm{DL}} + \lambda_i^{\mathrm{scan}}))Y_{\mathrm{AP}}}. \tag{4.5}$$

In the above-mentioned equations, $\lambda_i^{\mathrm{scan}}$ and $\lambda_i^{\mathrm{ACKscan}}$ are the scan-packet arrival rate to an AP for device i and the ACK response of the SPs from device i per unit time. Using the above-mentioned queue model and carrier-sense multiple-access/collision avoidance (CSMA/CA) process, we formulate the IoT and scan throughput over the IEEE 802.11ah RAW mechanism in the next subsection.

4.4.1.2 IoT/Scan Throughput Estimation

Before modeling the CSMA/CA process at per Rslot by a group of devices in IEEE 802.11ah, we estimate the number of groups per AP (M) and number of devices (g) per group. M is estimated based on the length of the RFrame and the available the Rslot, formulated as $\lceil t_{\mathrm{R}}/t'_{\mathrm{s}} \rceil$. Without loss of generality, we consider a uniform distribution of devices among M groups in this study. The number of devices per group (g) is equal to N/M. In the case of N not being divisible by M, g in the $M-1$ group is $\lfloor N/M \rfloor$, and the last group has $N-(M \times \lfloor N/M \rfloor)$ devices. During the CSMA/CA process, a device invokes a backoff process after a packet arrives in the queue. After the backoff, a device j ($1 \leq j \leq g$), from any group, first senses the idle channel during the distributed interframe space (DIFS) equal to $DIFS$ interval. If the channel is found idle after a DIFS, the device transmits the data followed by a short interframe space (SIFS). After the SIFS, an ACK is transmitted by the AP to indicate a successful transmission. The device again waits for a DIFS before initiating a backoff timer after data transmission. Each device initiates separate consecutive packet transmissions with a random backoff timer. In CSMA/CA, each device follows an exponential backoff that is selected in the range of (0, W-1), where W is the contention window size equal to $2^m W_{\min}$, depending on the re-transmission number (m). $W_{\min}$ is the contention window size of the first attempt. Consecutive mini-slots having lengths of ϕ are assigned for backoff counting within an Rslot, which are equal to the time required by a device to detect a packet transmitted by other devices.

In this study, we consider unsaturated traffic because of our queuing model, where there is a probability of having an empty queue. Moreover, we assume that the channel condition is ideal without any communication errors or capture effects. However, multiple transmission attempts in the same mini-slot can cause a collision and initiate a backoff. A device or an AP can transmit a packet in a

mini-slot when the queue has at least a packet to transmit. Thus, the transmission probability by device j and AP is given as follows:

$$\tau_j = \tau_{\text{sat}} Q_j^{\text{Dnempslot}} = \tau_{\text{sat}}(1 - e^{-(\lambda_j^{\text{UL}} t_{\text{s}} + \lambda_j^{\text{ACKscanslot}}) Y_j}), \tag{4.6}$$

$$\tau_{\text{AP}} = \tau_{\text{sat}} Q^{\text{APnempslot}} = \tau_{\text{sat}}(1 - e^{-(\sum_{j=1}^{g}(\lambda_j^{\text{DL}} t_{\text{s}} + \lambda_j^{\text{scanslot}})) Y_{\text{AP}}}), \tag{4.7}$$

where τ_{sat} is the saturated transmission probability of a terminal when the terminal and the AP always have a packet to transmit in a mini-slot. In Equations (4.6) and (4.7), $Q_j^{\text{Dnempslot}}$ and $Q^{\text{APnempslot}}$ are the probabilities of having at least a packet in the queue of device j and the AP during Y_i and Y_{AP}, respectively, derived based on (4.4) and (4.5). However, devices transmit in their Rslot. Hence, λ_j^{UL} and λ_j^{DL} are multiplied by t_{s} to estimate the average number of packets arriving per Rslot in (4.6) and (4.7). $\lambda_j^{\text{ACKscanslot}}$ and $\lambda_j^{\text{scanslot}}$ are the average numbers of scan and ACK response packets per Rslot that will be formulated later in the model. τ_{sat}, based on [45], is formulated as

$$\tau_{\text{sat}} = \frac{2(1 - C_{\text{sat}})}{(1 - 2C_{\text{sat}})(W_{\min} + 1) + C_{\text{sat}} W_{\min}(1 - (2C_{\text{sat}})^m)}, \tag{4.8}$$

where C_{sat} is the conditional collision probability of a packet when at least one of the remaining $g - 1$ devices or the AP transmits data. Hence, C_{sat} should be expressed as

$$C_{\text{sat}} = 1 - (1 - \tau_{\text{sat}})^{g-1}(1 - \tau_{\text{sat}}) = 1 - (1 - \tau_{\text{sat}})^g. \tag{4.9}$$

Equations (4.8) and (4.9) can be solved using numerical techniques. However, to estimate the IoT/scan throughput, we must formulate the successful transmission of IoT and SPs. Hence, P_j^{succ} and $P_{\text{AP}}^{\text{succ}}$ are the probabilities for successful

transmission by device j or the AP. A successful packet transmission occurs when exactly one device or one AP transmits on the channel when, at least one device/AP transmits during the considered mini-slot time. Based on this definition, P_j^{succ} and $P_{\text{AP}}^{\text{succ}}$ can be derived as

$$P_j^{\text{succ}} = \frac{(g+1)\tau_j(1-\tau_{\text{AP}})(1-\prod_{q=1,q\neq j}^{g}(1-\tau_q))}{1-\prod_{q=1}^{g}(1-\tau_q)(1-\tau_{\text{AP}})}. \tag{4.10}$$

$$P_{\text{AP}}^{\text{succ}} = \frac{(g+1)\tau_{\text{AP}}(1-\prod_{j=1}^{g}(1-\tau_j))}{1-\prod_{j=1}^{g}(1-\tau_j)(1-\tau_{\text{AP}})}. \tag{4.11}$$

q is a device-index term ranges from 1 to g. Before formulating the throughput, we must derive the equations for the service time in the queue of each device and the AP to substitute into (4.6) and (4.7). Here, the service time is defined in terms of the packet transmission time for each device or AP. The service time of a packet in the CSMA/CA process usually includes the waiting time during backoff and the transmission time. Moreover, a device cannot transmit packets during a sleep time. Thus, an unsuccessful packet in an Rslot must wait until the next Rslot for transmission (i.e., the sleep duration). Therefore, the service time for packets over IEEE 802.11ah should include the sleep time. The service time can thus be derived as

$$Y_j = T_j^{\text{bo}} N^{\text{bo}} + T_j^{\text{tr}} + (1-P_j^{\text{succ}})T_j^{\text{sleep}}, \tag{4.12}$$

where T_j^{bo} and N^{bo} are the mean length of a backoff mini-slot and the average number of backoff mini-slots, respectively, and T_j^{tr} and $T_j^{\text{sleep}} = t_{\text{R}} - t_{\text{s}}$ are the successful transmission time and sleep duration, respectively. Furthermore, T_j^{bo} is calculated on the basis of the empty-slot time with no transmission by $g-1$ devices or the AP, the successful transfer time by device j, and the idle time

during the collision:

$$T_j^{\text{bo}} = (1 - P^{\text{tr}})\phi + T^{\text{succ}} P^{\text{tr}} P_j^{\text{succ}} + T^{\text{col}}(1 - P_j^{\text{succ}})P^{\text{tr}}. \tag{4.13}$$

In (4.13), P^{tr} is the probability that at least one transmission is performed by a device or an AP in the given mini-slots. Here, T^{succ} and T^{col} represent the average time spent during successful transmission plus the collision of a device or an AP using basic access, respectively. Furthermore, P^{tr}, T^{succ}, and T^{col} are formulated as

$$P^{\text{tr}} = 1 - \prod_{j=1}^{g}(1 - \tau_j)(1 - \tau_{\text{AP}}). \tag{4.14}$$

$$T^{\text{succ}} = T^{\text{data}} + SIFS + 2\delta + T_{\text{ACK}} + DIFS, \tag{4.15}$$

$$T^{\text{col}} = T^{\text{data}} + DIFS + \delta, \tag{4.16}$$

where $T^{\text{data}} = H/datarate$ is the transmission of a data frame. H is the IEEE 802.11ah frame size whose size is the sum of frame header (14 bytes), MAC payload, and frame control sequence (4 bytes) [46]. MAC payload size is the sum of IPv6 header size (P_{IPv6}), IPv6 extension header size (E_{IPv6}) and P, where P is the payload size of IoT data. IPv6 extension header is used to specify the ESP header. Furthermore, δ is the propagation delay, and $T_{\text{ACK}} = ACK/datarate$ is the transmission time of the ACK frame of size ACK. In addition, T_j^{tr} in (4.12) can be formulated as

$$T_j^{\text{tr}} = P_j^{\text{succ}} T^{\text{succ}} + (1 - P_j^{\text{succ}})T^{\text{col}}. \tag{4.17}$$

For simplicity, the constant, N^{bo}, in (4.12) is considered to be $W/2$ in this study. However, the average number of backoffs based on an exponential distribution

function can also be considered here as in other studies [45]. Similarly, we can formulate the service time for an AP as follows:

$$Y_{\mathrm{AP}} = T_{\mathrm{AP}}^{\mathrm{bo}} N^{\mathrm{bo}} + T_{\mathrm{AP}}^{\mathrm{tr}}, \tag{4.18}$$

$$T_{\mathrm{AP}}^{\mathrm{bo}} = (1 - P^{\mathrm{tr}})\phi + T^{\mathrm{succ}} P^{\mathrm{tr}} P_{\mathrm{AP}}^{\mathrm{succ}} + T^{\mathrm{col}}(1 - P_{\mathrm{AP}}^{\mathrm{succ}})P^{\mathrm{tr}}. \tag{4.19}$$

$$T_{\mathrm{AP}}^{\mathrm{tr}} = P_{\mathrm{AP}}^{\mathrm{succ}} T^{\mathrm{succ}} + (1 - P_{\mathrm{AP}}^{\mathrm{succ}})T^{\mathrm{col}}. \tag{4.20}$$

For $Q_j^{\mathrm{Dnempslot}}$ and $Q^{\mathrm{APnempslot}}$ of (4.6) and (4.7), we must estimate $\lambda_j^{\mathrm{scanslot}}$ and $\lambda_j^{\mathrm{ACKscanslot}}$ per Rslot, respectively, before formulating the IoT/scan throughput. The number of scan packets arriving at each AP (R) should be equal to TR/N_{AP}. The AP distributes these SPs according to the distribution of devices in groups. Because we considered the uniform distribution of devices, the AP uniformly distributes the SPs among the groups in each RFrame. Thus, $R' = R/(K \times M)$ is the number of SPs per Rslot. However, the distribution of SPs to each IoT device of a group in the Rslot is unknown, and it was not formulated in previous studies. Therefore, we propose a probability mass function (PMF) (P_{scan}) for the distribution of SPs per slot to a device in the group. Out of R', any number of packets can be assigned to a device. The derivation of PMF is presented in the APPENDIX and is expressed as follows:

$$P_{\mathrm{scan}}(X = x) = \frac{{}^{g-x+R'-2}\mathbf{C}_{R'-2}}{{}^{g+R'-1}\mathbf{C}_{R'-1}}, \tag{4.21}$$

where ${}^{g+R'-1}\mathbf{C}_{R'-1}$ is the total number of ways SPs can be distributed among g devices, and ${}^{g-x+R'-2}\mathbf{C}_{R'-2}$ is number of ways x SPs can target a device. The

average number of SPs aiming for device j in an Rslot can be formulated as

$$\lambda_j^{\text{scanslot}}(g,R) = \sum_{x=1}^{R'} x \times P_{\text{scan}}(X=x), \tag{4.22}$$

where R' is the function of R. The average number of ACK SPs ($\lambda_j^{\text{ACKscanslot}}$) sent by device j per Rslot can be expressed as

$$\lambda_j^{\text{ACKscanslot}}(g,R) = P_{\text{sat}}^{\text{succ}} \lambda_j^{\text{scanslot}}, \tag{4.23}$$

where $P_{\text{sat}}^{\text{succ}}$ is the success probability in the saturation case. The saturated probability is used to estimate the number of ACKs/RSTs, because these packets are received only if SPs are transmitted. Hence, $P_{\text{sat}}^{\text{succ}}$ should be used to estimate successful scan-packet transmission, and it can be derived from [9].

$$P_{\text{sat}}^{\text{succ}} = \frac{(g+1)(1-\tau_{\text{sat}})^g}{1-(1-\tau_{\text{sat}})^{(g+1)}}. \tag{4.24}$$

In (4.24), 1 is added for the AP in the expression (g+1). The IoT (Th_{IoT}) and scan (Th_{scan}) throughput can be derived using (4.4)–(4.24). The formulation of Th_{scan} is given by

$$Th_{\text{scan}} = \sum_{j=1}^{g} (P_{\text{AP}}^{\text{succ}} \lambda_j^{\text{scanslot}} + P_{\text{j}}^{\text{succ}} \lambda_j^{\text{ACKscanslot}}) L_{\text{scan}} MK, \tag{4.25}$$

where L_{scan} is the packet size of the SPs. Similarly, the IoT throughput can be expressed as

$$Th_{\text{IoT}} = \sum_{j=1}^{g} (\lambda_j^{\text{UL}} P_j^{\text{succ}} + \lambda_j^{\text{DL}} P_{\text{AP}}^{\text{succ}}) L_{\text{IoT}}, \tag{4.26}$$

where L_{IoT} is the size of the UL and DL packets. Without a loss of generality, we consider the same packet size for both UL and DL data. Based on the IoT

and scan throughputs, we present the security model to estimate the weights of CONF, INT, AVA, and attack likelihood. Finally, we formulate the risk as an objective function.

4.4.2 Risk Evaluation Model

In this subsection, we provide the complete mathematical formulation for risk, which is function of the CVSS value, as expressed in (4.1). Owing to changes in the temporal and environmental metrics, the overall CVSS score is given by [42]:

$$V = \text{round}(\text{round}(\min(6.42 \times MISS + Modexp), 10) \times Temp_{\text{metric}}), \tag{4.27}$$

$$MISS = \min(1 - ((1 - C_{\text{req}} \times \boldsymbol{C}_{\mathbf{mod}})(1 - I_{\text{req}} \times \boldsymbol{I}_{\mathbf{mod}}) (1 - A_{\text{req}} \times \boldsymbol{A}_{\mathbf{mod}})), 0.915), \tag{4.28}$$

$$Temp_{\text{metric}} = \boldsymbol{AL} \times RL \times RC. \tag{4.29}$$

In (4.27), $MISS$ is the modified impact score denoting the effects of attacks on any system causing changes in CONF, INT, and AVA, compared with their requirements, as defined in (4.28). In (4.28), C_{req} is the confidentiality requirement that is set according to the type of data (e.g., sensitive) communicated by an application or a system. I_{req} is the integrity requirement that depends on the importance of the accuracy of data communicated by an application, including healthcare data, which has steep requirements. Similarly, A_{req} is the availability requirement that is based on the importance of the accessibility of resources in the system. From (4.27) and (4.28), V is a function of the modified CONF (C_{mod}), modified INT (I_{mod}), and modified AVA (A_{mod}), modified because of changes in

the environment, including available network resources. Moreover, V is also a function of the $Temp_{\text{metric}}$, which is function of attack likelihood (AL), report confidence (RC), and remediation level (RL). AL varies because of changes in patching delay ($Pdelay$), which is formulated later in this section. The parameter RL and RC are independent of scan rate. Hence, we do not formulate them here. However, constant values are given as parameter settings in the next section. The details and importance of each parameter can be found in [42]. Furthermore, $Modexp$ is the modified exploitability that reflects changes in attack approaches by attackers and characteristics of vulnerable components. Its expression is taken from [42].

$$\begin{aligned} Modexp =& 8.22 \times ModifiedAttackVector \\ & \times ModifiedAttackComplexity \\ & \times ModifiedPrivilegesRequired \\ & \times ModifiedUserInteraction. \end{aligned} \tag{4.30}$$

In (4.30), $ModifiedAttackVector$ reflects a change in attack position of an attacker from any network, remotely, in an adjacent network, locally, etc. $ModifiedAttackComplexity$ describes the modification in attacker capability or skill to exploit a vulnerability with information about the complexity of their approach. Moreover, $ModifiedPrivilegesRequired$ represents a change in an attacker's privilege before exploiting a vulnerability. Similarly, $ModifiedUserInteraction$ includes the need for any human access other than an attacker's to compromise the target network. The parameters in (4.30) are independent of the scan rate and depend on attackers' skills and approaches. Hence, we consider the value from [42], which is discussed in the Results section.

CONF and INT of IoT packets are provided by IPsec from the encryption and authentication of IoT packets transmitted between devices over untrusted

networks. Moreover, we can say that C_{mod} and I_{mod} of (4.28) depend on the strength of the encryption algorithm. With IPsec, the sec-admin defines a set of encryption algorithms for devices based on the observation of traffic and QoS requirements, as discussed in Section 4.2. Devices negotiate and choose an encryption algorithm that is transmitted during a security association. In this way, the ESP protocol gets the information about the encryption algorithm to be used. To generalize the model, we consider A to be the number of algorithms suitable for IoT. These algorithms can be identified as Algo_E, where E varies from 1 to A. In this study, we assign a weight value (E) to these algorithms according to their strengths (A). Thus, Algo_A has $E = A$ weight, which has a higher strength than the weight of Algo_{A-1} (i.e., $E = A - 1$), $A - 2$,....., and 1. As discussed, sec-admins define the encryption algorithm for IoT devices based on QoS requirements (e.g., throughput and delay) [37,39]. Moreover, sec-admins can decide not to use any encryption algorithm or a weaker one in case of poor performance to meet the QoS requirements. Thus, we include no encryption for this thesis. With this model, we represent the assigned encryption algorithm using a weight value (WE) for each device on the basis of the throughput ratio (TH_{ratio}) (i.e., $Th_{\text{IoT}}/(Th_{\text{IoT}} + Th_{\text{scan}})$). The weight based on the throughput ratio is defined as

$$WE = E, \text{ if } \frac{E-1}{A} < TH_{\text{ratio}} < \frac{E}{A}, E \in [1, A], \mathbb{N}. \tag{4.31}$$

For example, if TH_{ratio}=0.15 and A=5, then E can be calculated on the basis of the condition in (4.31), which is equal to two. From (4.31), WE is a function of IoT throughput, which is a function of scan rate. The weights of C_{mod} and I_{mod} are determined according to the ratio of the assigned WE to their maximum

weight (A). However, C_{mod} and I_{mod} are defined as

$$C_{\text{mod}} = \frac{WE}{A}, \tag{4.32}$$

$$I_{\text{mod}} = \frac{WE}{A}. \tag{4.33}$$

C_{mod} and I_{mod} have the same formulae, owing to the ESP protocol with authentication, which provides both CONF and INT by encrypting payload and authentication data using the same encryption algorithm. Furthermore, we must estimate A_{mod}. Hence, A_{mod} represents the impact of the availability on the IoT, which can be expressed as the number of unsuccessful transmissions of IoT packets out of the total number of packets sent by a device in an Rslot. Hence, A_{mod} for a device, j, can be expressed as

$$A_{\text{mod}} = (1 - P_j^{\text{succ}})(\frac{\lambda_j^{\text{UL}} t_{\text{s}}}{\lambda_j^{\text{UL}} t_{\text{s}} + \lambda_j^{\text{ACKscanslot}}}). \tag{4.34}$$

After formulating the environmental sub-metrics, we derive the expression of our focused temporal sub-metric (AL), which depends on the patching delay (PD) of a device. AL is also function of mean time to compromise (MTTC) any device, denoted as TC. MTTC is estimated by an attacker ' s expertise and tools. However, there have been many studies that have estimated MTTC duration. An attacker usually takes a few days to attack any system based on abilities and tools. However, evaluating MTTC is beyond the scope of this study. We instead consider a constant value that can be substituted from [47]. Thus, AL can be defined as the ratio of PD to MTTC. However, if the MTTC is less than the patching delay, then there will be a 100% likelihood that the device will

be attacked. Based on this definition, AL can be formulated as

$$AL = \begin{cases} \frac{Pdelay}{TC} & \text{if } Pdelay < TC, \\ 1 & \text{Otherwise.} \end{cases} \tag{4.35}$$

The patching delay is defined as the time required to fix the existing vulnerabilities in a device. There are two steps needed to patch the vulnerabilities: identifying vulnerabilities in any port of a device and implementing patching algorithms. Thus, the patching delay is equal to the time taken to identify vulnerabilities and the patching algorithm processing time. Hence, we scan all ports of a device to discover the vulnerabilities. The scan delay (t_{scan}) includes the identification time for all the vulnerabilities of a device. Moreover, we assume that patching algorithms are available. Hence, the processing time (t_{patch}) for patching is constant. The patching delay is formulated as

$$Pdelay = t_{\text{scan}} + t_{\text{patch}}. \tag{4.36}$$

In this study, we suppose that a SP is dedicated to probing a device port. Moreover, the port scanner performs a horizontal scan, thus scanning all the ports. However, as explained in Section 4.3, the scan delay varies with the scan rate and network performance of the WLAN. Thus, the scan delay can be formulated based on

$$t_{\text{scan}} = \frac{N_{\text{P}} t_b}{N_{\text{Port}}^{\text{succ}}}, \tag{4.37}$$

where N_{P} and $N_{\text{Port}}^{\text{succ}}$ are the total number of ports per device required to be probed and the number of successfully scanned ports per beacon interval, respectively. $N_{\text{Port}}^{\text{succ}}$ is calculated by the number of successful SYN packet transmissions and ACK packets per beacon interval, except for waiting packets in their queue,

expressed as

$$N_{\text{Port}}^{\text{succ}} = (P_{\text{AP}}^{\text{succ}}(\lambda_j^{\text{scanslot}} - R_{\text{AP}}^{\text{Qscan}}) - R_j^{\text{QACKscan}})P_{\text{avg}}^{\text{succ}}MK. \tag{4.38}$$

$\lambda_j^{\text{scanslot}}$ is defined in (4.22). $P_{\text{avg}}^{\text{succ}}$ is the average probability of successfully transmitting a packet by any device within a group, formulated as

$$P_{\text{avg}}^{\text{succ}} = \frac{\sum_{j=1}^{g} P_j^{\text{succ}}}{g}. \tag{4.39}$$

In (4.38), $R_{\text{AP}}^{\text{Qscan}}$ and R_j^{QACKscan} are the average numbers of scan and ACK packets waiting in the queue of the AP and any device per Rslot, respectively. In this study, we consider a stable and infinite buffer (i.e., $\lambda < \frac{1}{Y}$). $R_{\text{AP}}^{\text{Qscan}}$ and R_j^{QACKscan} can be formulated using Little's theorem, expressed as

$$R_j^{\text{QACKscan}} = \frac{(\lambda_j^{\text{ACKscan}})^2}{\frac{1}{Y_j}(\frac{1}{Y_j} - \lambda_j^{\text{ACKscan}})}. \tag{4.40}$$

$$R_{\text{AP}}^{\text{Qscan}} = \frac{(\lambda_j^{\text{scan}})^2}{\frac{1}{Y_{\text{AP}}}(\frac{1}{Y_{\text{AP}}} - \lambda_j^{\text{scan}})}. \tag{4.41}$$

The CVSS score of any vulnerability can be derived using (4.31)–(4.41) in the CVSS equation, as presented in (4.27). The value of the calculated V of a device is then substituted in (4.1) to finally get the objective function to minimize the risk. The final optimization equation is given by

$$\begin{aligned} &\underset{TR}{\text{Minimize}} \quad Risk(TR), \\ &\text{Subject to:} \quad TR, t_{\text{b}}, t_{\text{R}}, t_{\text{s}}, N > 0. \end{aligned} \tag{4.42}$$

From the overall mathematical model, we observe that the scan rate (R) influences the IoT throughput and scan delay, which, in turn, degrades encryption algorithms and increases the patching delay, respectively. Hence, we solve for the

optimal R in the next section using numerical analysis and present our results, which show that the risk to any IoT device can be minimized.

4.5 Numerical Analysis

Table 4.1: Parameter settings.

Variables	*Values*
λ_j^{ULPU} (packets/second) λ_j^{DLPU},λ_B^{UL}, λ_B^{DL}	j, $j \in [1, 10]$ 2, 5, 5
N_{RAW}, t_b, t_s	2, 1 s, 50 ms
$DIFS$, $SIFS$, ϕ [9]	252 μs, 160 μs, 52 μs, 0.132 ms
$W_{\min}$, m [9]	16, 7
$P = L_{\text{scan/IoT}}$, ACK [9]	64 bytes, 250 bits
IPv6 parameters: P_{IPv6}, E_{IPv6}	40 bytes, 2 bytes
$datarate$, δ	4 Mbps, 1μs
Th, AI, RC, RL, TC [41] [47]	1, 1, 0.95, 1, 10 days

4.6 Performance Evaluation

This section presents a numerical analysis of the mathematical model proposed in the previous section. According to the model, we initially analyze the IoT throughput at various scan rates per AP, which leads to changes in the weights of CONF, INT, and AVA. Furthermore, we show the effect of variations in the attack likelihood and patching delay on the scan rates per AP. In this analysis, our objective is to minimize the risk to an IoT device at an optimal scan rate.

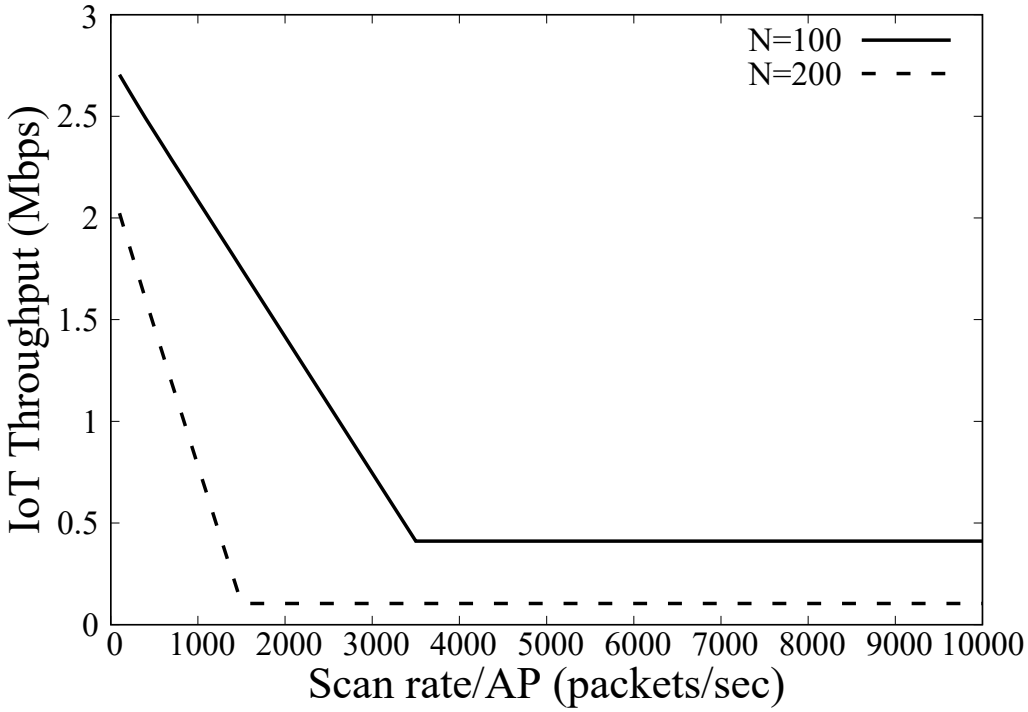

Figure 4.2: IoT throughput at various scan rate.

Thus, we finally analyze the objective function. Doing so, we can vary only N and R, because sec-admins are oblivious to other network settings. In this analysis, IEEE 802.11ah frame and IPv6 packet structure parameters are used as formulated in section 4.4.1.2, which are defined in Table 4.1 with other required parameters for the analysis.

4.6.1 IoT Throughput and Environmental Metrics Analysis

Figure 4.2 shows the evaluation of the total IoT data throughput at various scan rates for $N = 100, 200$ Het-IoT devices per AP over the IEEE 802.11ah RAW mechanism. This analysis is performed on the basis of (4.26), where $P_{\mathrm{AP}}^{\mathrm{succ}}$ and

$P_{\mathrm{j}}^{\mathrm{succ}}$ are derived from (4.11) and (4.10). From the result, we can observe that the IoT throughput decreases as the scan rate increases, ultimately reaching a saturation state (i.e., constant line). The reason for such behavior is the availability of channel resources at a low scan rate. Moreover, a low scan rate does not contribute significantly to network congestion via IoT packets. Furthermore, the rationale for saturated IoT throughput after a certain scan rate is the achievement of the maximum system load. We can also observe that $N = 200$ achieves a higher IoT throughput than does $N = 100$, owing to channel access by a greater number of devices per group for the former, which congests the limited network per slot. According to the IoT throughput, we determine the modified CONF and INT in Fig. 4.3. For this analysis, we consider five encryption algorithms that are suitable for IoT-type devices. Hence, A is equal to five in (4.31). The encryption algorithms include no encryption, DES-HW, DES, 3DES-HW, and AES-128, in increasing order of strength. We analyze C_{mod} and I_{mod} from (4.32) and (4.33), which are functions of TH_{ratio} according to WE. As discussed, the IoT throughput decreases as the scan rate increases, which forces security admins to set weaker algorithms or no encryption at a high scan rate. Consequently, the values of WE, C_{mod}, and I_{mod} decrease greatly. However, C_{mod} and I_{mod} are constant after a certain scan rate, owing to the saturated IoT throughput. For N=200, the weights of C_{mod} and I_{mod} are lower than those for $N = 100$, owing to the lower IoT throughput of the former. Similarly, analysis of another parameter of environmental metrics (i.e., A_{mod}) is shown in Fig. 4.4, which is evaluated using (4.34). In Fig. 4.4, we can see that the availability of packets decreases as the scan rates increase, owing to an increase in the collision of IoT packets with a large number of SPs. The availability of packets is higher for N=100 compared with N=200, as in the former case, because there are few devices/groups. Thus, the congestion in every Rslot is reduced. This impact of SPs on regular data can also

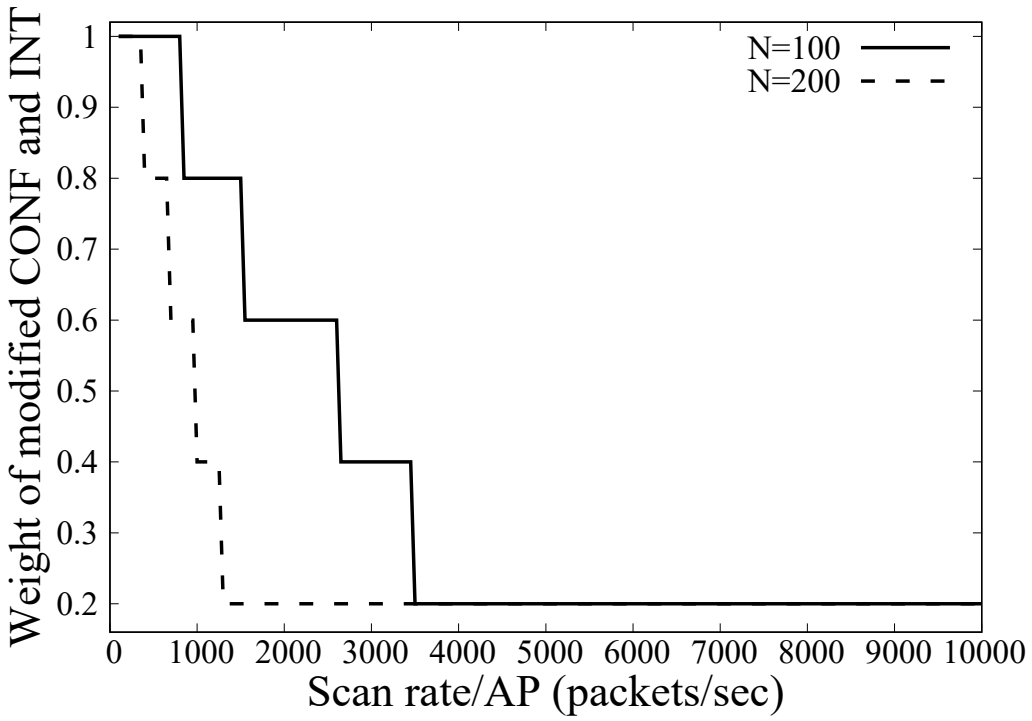

Figure 4.3: Variation of C_{mod} and I_{mod} at various scan rates.

be considered to be a scan attack [11]. Hence, our approach to optimize the scan rate can avoid such types of attacks as well. From the discussion presented above, we can conclude that an appropriate scan rate should be set such that environmental sub-metrics can lead to high performance without significantly affecting the security.

4.6.2 Temporal Metrics Analysis

The analysis of the temporal metrics is shown in Figs. 4.5 and 4.6. From Fig. 4.5, we observe that the patching delay is high at low scan rates, owing to the slow scan performance. It decreases to a certain scan rate. Thereafter, it increases because of network congestion and unsuccessful transmission of SPs. Unsuccessful transmission is responsible for the increase in the scan delay of the identification

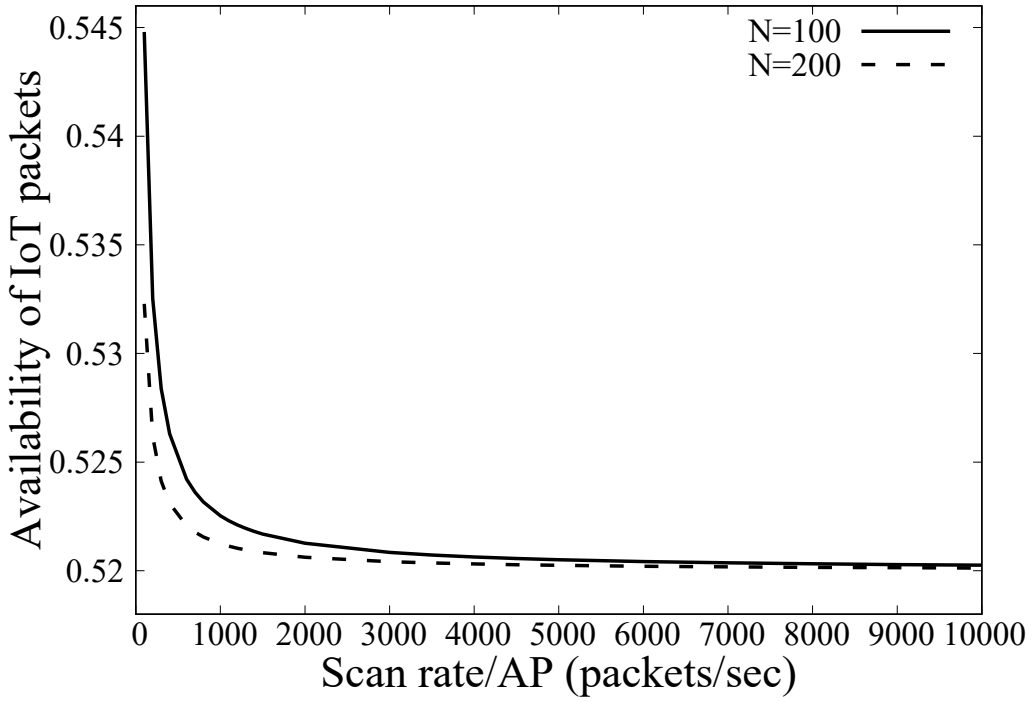

Figure 4.4: Availability of IoT packets at different scan rates.

of vulnerabilities or services at all ports. Attack likelihood is a function of the patching delay, as determined by (4.35). Hence, in Fig. 4.6, the attack likelihood has a higher value at low or high scan rates, and it is directly proportional to the patching delay. Therefore, it is imperative to set an appropriate scan rate such that the attack likelihood is minimized. Moreover, the patching delay and attack likelihood for N=100 are lower than those for N=200, owing to the small number of devices per group, which reduces the congestion at each Rslot. From this discussion, we can conclude that an appropriate scan rate for N IoT devices will reduce the impact on the temporal metrics of CVSS, which in turn will improve the security.

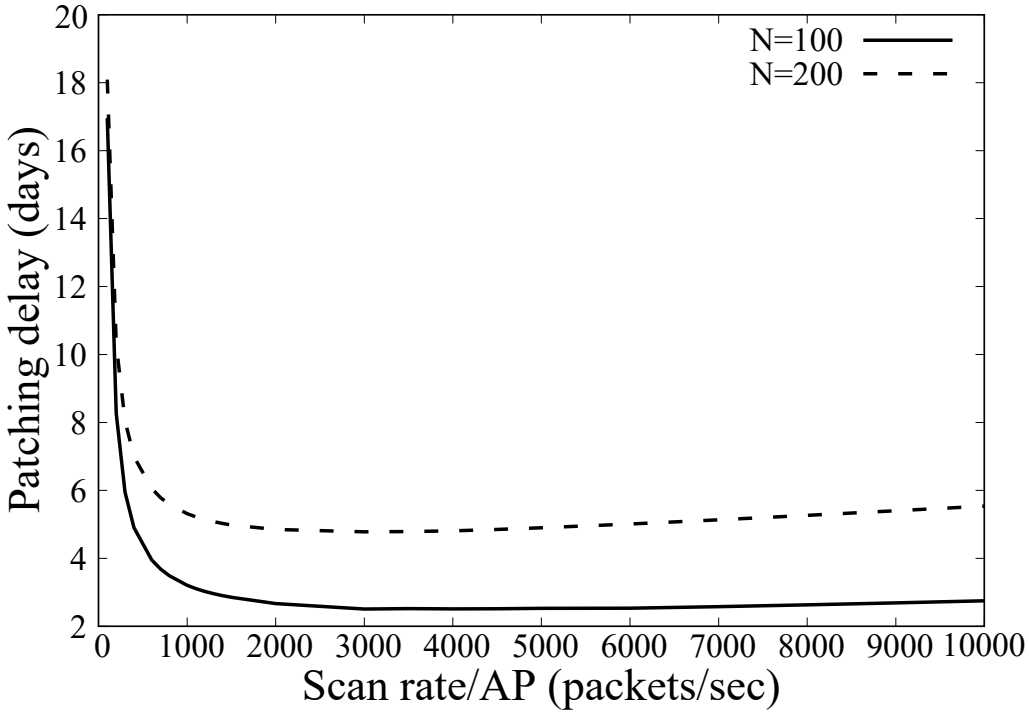

Figure 4.5: Patching-delay analysis at various scan rates.

4.6.3 Risk Minimization

Our objective in this analysis is to minimize the risk caused by the impact of the scan rate on the security services of an IoT device. Finally, based on the previous analysis and risk formulae, we present the result of risk evaluation at various scan rates in Fig. 4.7. We can observe that the risk value is high at lower scan rates owing to the influence of a high patching delay and high attack likelihood. For instance, the risk value of 8.3 at a lower scan rate=100 for N=100. As quantitative values of risk provided in Section 4.2, maximum value of risk can be 10, that denotes low security of an IoT device. The risk value of 8.3 is close to 10, which indicates that the risk on a IoT device is higher at a lower scan rate. Similarly, the risk value at a higher scan rate such as 9000 for N=200 is 9.95, which is very close to 10. This implies that the IoT device can be highly

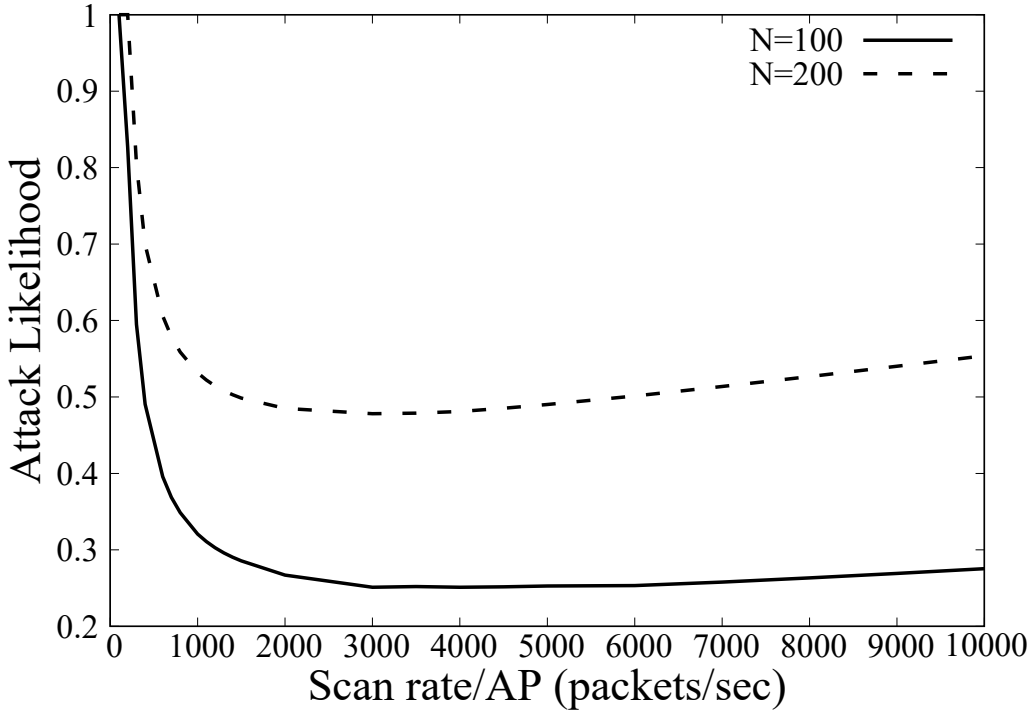

Figure 4.6: Attack likelihood for an IoT device.

unsecured. Therefore, the risk is minimized for an IoT device to 1.91 in the case of N=100 at R=3,500. The risk value 1.91 is close to 0, which means it has low risk. Similarly, for N=200, the risk is minimized at R=1,300 and obtained risk value is 4.6, which means risk is medium in this case. The optimal scan rate is high for N=100 than N=200 owing to low congestion at network due to less number of devices per group in N=100. From Fig. 4.7 and quantitative analysis, we can see that our proposed model provides an optimal scan rate that minimizes risk and improves the security of IoT device.

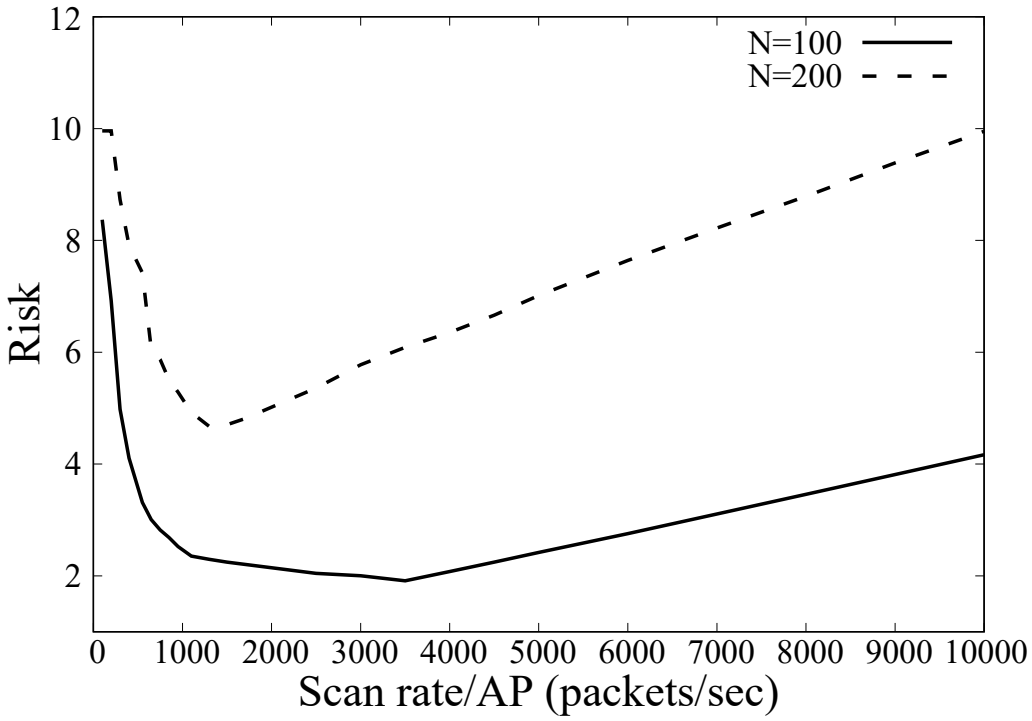

Figure 4.7: Risk to each IoT device.

4.7 Summary

In this chapter, We proposed a novel network-aware IWPS used to set an optimal scan rate for sec-admins such that the security of IoT devices is maximized over emerging WLANs, such as IEEE 802.11ah. To include heterogeneous traffic of IoT and scan-packet arrival, we designed a novel queue model based on the Markov chain for each IoT device and AP. Then, we proposed a mathematical model to estimate the IoT and scan throughput based on the IEEE 802.11ah RAW mechanism, the CSMA/CA process, and the scan traffic for the proposed queue model. We also formulated a mathematical model used to evaluate the risk to each IoT device, which consists of new models used to assess CONF, INT, AVA, and AL at various scan rates. Using these models, we performed numerical analyses

to ascertain the effect of the scan rate on the variations of CONF, INT, AVA, and AL caused by low network performance (e.g., IoT throughput, unsuccessful IoT packet transmissions, and scan delays) under scan-rate variations. We optimized this tradeoff by minimizing the risk to each device at an optimal scan rate in the final analysis. From our analyses, we observed that an optimal scan rate provides high security while ensuring QoS.

Chapter 5

Conclusion

The realization of smart and innovative services such as healthcare, intelligent transport system, and so on depends on the efficient IoT system. The key elements of IoT is the connectivity which plays major role in the realization of such services. The key requirements of IoT wireless network are massive connection, energy efficiency, and secure communication. In this vein, standards have been developing communication technologies for IoT wireless network. 5G has been developed and included new features for IoT-WWAN and IEEE 802.11ah is proposed for IoT-WLAN. However, both communication technologies have tradeoff in terms of meeting requirements. The LTE-A pro networks under 5G has technology such as group paging to enable massive connection and have other technologies for end-to-end encryption for security. However, group paging is not energy efficient. Similarly, IEEE 802.11ah has features like novel MAC protocol for energy efficient and massive communications but still carries existing security issues. Therefore, in this thesis, we focused on design energy efficient group paging for IoT over LTE-A Pro networks under 5G, and IWPS for securing IEEE 802.11ah enabled IoT-WLAN. Overall, this thesis designed models for energy efficient and secure data communication system for IoT wireless network. Precisely,

our contributions are listed below: In chapter 2, the overview of data communication technologies over LTE-A Pro networks that is group paging was introduced. Additionally, we mentioned the existing related works of group paging with novel research direction. We have also introduced the overview IEEE802.11ah MAC protocols and IWPS system for vulnerability identification in IoT-WLAN devices. The related works of IWPS was also presented in this chapter.

In chapter 3, we discussed the research challenges related to group paging that is energy efficient grouping of MIDs while providing packet delay and loss requirements. In this chapter, we proposed novel mathematical models to find the optimal groups for MIDs. In this model, we formulated novel packet arrival delay and packet loss model for group paging system. Thereafter, we introduced new energy models for the considered system. Based on these models, we formulated convex optimization problem and provided the solution using Lagrangian approach and Karush-Kuhn-Tucker conditions. Finally, we have analyzed the models for groups having various characteristics and also compared the results with conventional random grouping approach. Based on analysis, this chapter concluded that our proposed solution saves significant amount of energy.

In chapter 4, we identified novel research problem of IWPS that can degrade the IoT security by scanning IoT devices without the knowledge of network. To solve such complex issue, in this chapter, we proposed novel mathematical models to design network-aware IWPS for maximization of IoT security. First, we proposed IoT traffic and network models to estimate IoT throughput in the presence of scan packets at the WLAN. Based on these models, we modeled IoT security by formulating risk metric as function of scan rate, and then calculated the optimize scan rate to maximize IoT security and minimize risk. This chapter also presented the numerical analysis of proposed models and confirmed that network-oblivious scan can degrade the IoT security in spite of improving it.

In chapter 5, we concluded this thesis.

As mentioned above, in this research, we have improved a new feature of LTE-A pro network under 5G and improving the IWPS for maximizing security. Hence, our proposed models and results will be big contribution for meeting the requirements of IoT during the deployment of these technologies.

Appendix

In this Appendix, we present the derivation of the PMF of the distribution of SPs (P_{scan}) to a device in an Rslot (i.e., as stated in (4.21)). Let us assume that SPs arriving per Rslot are distributed randomly. Hence, we derive the distribution of incoming SPs per slot. Let X be a discrete random variable signifying the number of SPs assigned to a device, $x \in [1, R']$. There are g devices in a group. Hence, total number of ways SPs ($total_{\text{way}}$) can distribute among g devices can be given as formulation of distribution of identical objects into distinct bins (i.e., identical SPs to distinct devices), expressed below:

$$total_{\text{way}} = {}^{g+R'-1}\mathbf{C}_{R'-1}. \tag{A.1}$$

However, the number of ways SPs ($total_{\text{xway}}(x)$) can be distributed, given a device receives x SPs that range from 1 to R', is formulated below:

$$total_{\text{xway}}(0) = {}^{g+R'-2}\mathbf{C}_{R'-2}$$

$$total_{\text{xway}}(1) = {}^{g-1+R'-2}\mathbf{C}_{R'-2}$$

$$\vdots$$

$$total_{\text{xway}}(R') = {}^{g-R'+R'-2}\mathbf{C}_{R'-2}. \tag{A.2}$$

Based on (A.2), the expression for $total_{\text{xway}}(x)$ can be written as

$$total_{\text{xway}}(x) = {}^{g-x+R'-2}\mathbf{C}_{R'-2} \quad x \in [1, R']. \tag{A.3}$$

The PMF for receiving x packets by a device can be written as

$$P_{\text{scan}}(X = x) = \frac{{}^{g-x+R'-2}\mathbf{C}_{R'-2}}{{}^{g+R'-1}\mathbf{C}_{R'-1}}. \tag{A.4}$$

Copyright Permissions

We enclose the permissions that were used to write this dissertation. Please see the attached documents for a detailed description of the permissions. The publications used to write this thesis are listed as follows:

- Shikhar Verma, Y. Kawamoto and N. Kato, "A Network-aware Internet-wide Scan for Security Maximization of IPv6-enabled WLAN IoT Devices," in *IEEE Internet of Things Journal* (Accepted).

- Shikhar Verma, Y. Kawamoto and N. Kato, "Energy-Efficient Group Paging Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G," in *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 9187-9199, Oct. 2019.

- Shikhar Verma, Y. Kawamoto, and Nei Kato, "Security Analysis of Network-Oblivious Internet-Wide Scan for IEEE 802.11ah Enabled IoT," *IEEE International Conference on Vehicular Communications (VTC-Fall 2020), Virtual Conference,* Nov. 2020.

A Network-aware Internet-wide Scan for Security Maximization of IPv6-enabled WLAN IoT Devices

Author: Shikhar Verma
Publication: IEEE Internet of Things Journal
Publisher: IEEE
Date: Dec 31, 1969

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Energy-Efficient Group Paging Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G

Author: Shikhar Verma
Publication: IEEE Internet of Things Journal (J-IOT)
Publisher: IEEE
Date: Oct. 2019

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK CLOSE WINDOW

Publications

Journals

[1] Shikhar Verma, Y. Kawamoto and N. Kato, "A Network-aware Internet-wide Scan for Security Maximization of IPv6-enabled WLAN IoT Devices," in *IEEE Internet of Things Journal*, Early Access.

[2] Shikhar Verma, Y. Kawamoto and N. Kato, "Energy-Efficient Group Paging Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G," in *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 9187-9199, Oct. 2019.

[3] Shikhar Verma; Y. Kawamoto; Z. Fadlullah; H. Nishiyama; N. Kato, "A Survey on Network Methodologies for Real-Time Analytics of Massive IoT Data and Open Research Issues," in *IEEE Communications Surveys & Tutorials,* vol. 19, no. 3, pp. 1457-1477, April 2017.

Refereed Conference Papers

[4] Shikhar Verma, Y. Kawamoto, and Nei Kato, "A Novel IoT-Aware WLAN Environment Identification for Efficient Internet-Wide Port Scan," *IEEE International Global Communication Conference (GLOBECOM),* Taipei, Taiwan, Dec. 2020.

[5] Shikhar Verma, Y. Kawamoto, and Nei Kato, "Security Analysis of Network-Oblivious Internet-Wide Scan for IEEE 802.11ah Enabled IoT," *IEEE In-*

ternational Conference on Vehicular Communications (VTC-Fall 2020), Virtual Conference, Nov. 2020. (Student paper award)

[6] Shikhar Verma, Y. Kawamoto, H. Nishiyama, N. Kato and C. Huang, "Novel Group Paging Scheme for Improving Energy Efficiency of IoT Devices over LTE-A Pro Networks with QoS Considerations," *IEEE International Conference on Communications (ICC),* Kansas City, MO, May. 2018, pp. 1-6. (Best paper award)

Domestic Conference Papers

[7] Shikhar Verma, Y. Kawamoto, and Nei Kato, "A Study on WLAN environment Classification for Efficient Internet-wide Port Scan Design", IEICE General Conference-2021, Virtual, Mar. 2021.

[8] Shikhar Verma, Y. Kawamoto, and Nei Kato, "A Study on the Impact of Internet-wide Scan on QoS of IoT over IEEE 802.11ah", IEICE General Conference-2020, Hiroshima, Japan, Mar. 2020.

Awards

[9] IEEE Global Communications (GLOBECOM 2020), Student Travel Award, "A Novel IoT-Aware WLAN Environment Identification for Efficient Internet-Wide Port Scan," Taiwan, Dec. 2020

[10] IEEE International Conference on Vehicular Communications (VTC-Fall 2020), Student Paper Award, VTS Japan Society, "Security Analysis of Network-Oblivious Internet-Wide Scan for IEEE 802.11ah Enabled IoT," Tokyo, Japan, Nov. 20202.

[11] IEEE International Conference on Communications 2018, Best Paper Award, "Novel Group Paging Scheme for Improving Energy Efficiency of IoT Devices over LTE-A Pro Networks with QoS Considerations," Kansas, USA, May 2018.

[12] IEEE International Conference on Communications 2018, Student Travel Award, "Novel Group Paging Scheme for Improving Energy Efficiency of IoT Devices over LTE-A Pro Networks with QoS Considerations," Kansas, USA, May 2018.

[13] Research Excellence Award, IEEE Communication Society Sendai, 2018

References

[1] S. Verma, Y. Kawamoto, Z. M. Fadlullah, H. Nishiyama and N. Kato, "A Survey on Network Methodologies for Real-Time Analytics of Massive IoT Data and Open Research Issues," *IEEE Communications Surveys & Tutorials,* vol. 19, no. 3, pp. 1457-1477, Apr. 2017.

[2] IHS Market, "The Internet of Things: a movement, not a market", Available online: https://cdn.ihs.com/www/pdf/IoT_ebook.pdf (accessed on 5 Jan. 2020).

[3] 3rd Generation Partnership Project (3GPP), "Standars for the IoT", https://www.3gpp.org/news-events/1805-iot_r14 (accessed on 5 Jan. 2020)

[4] S. Aust, R. V. Prasad and I. G. M. M. Niemegeers, "IEEE 802.11ah: Advantages in standards and further challenges for sub 1 GHz Wi-Fi," *IEEE International Conference on Communications (ICC) 2012,* Ottawa, ON, Canada, pp. 6885-6889, Jun. 2012.

[5] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," in *IEEE Communications Surveys & Tutorials,* vol. 17, no. 4, pp. 2347-2376, Jun. 2015.

[6] S. Verma, Y. Kawamoto and N. Kato, "Energy-Efficient Group Paging Mechanism for QoS Constrained Mobile IoT Devices Over LTE-A Pro Networks Under 5G," in *IEEE Internet of Things Journal,* vol. 6, no. 5, pp. 9187-9199, Oct. 2019.

[7] N. Vlajic and D. Zhou, "IoT as a Land of Opportunity for DDoS Hackers," in *IEEE Computer,* vol. 51, no. 7, pp. 26-34, July 2018.

[8] O. Arouk, A. Ksentini and T. Taleb, "Group paging-based energy saving for massive MTC accesses in LTE and beyond networks", *IEEE Journal in Selected Areas of Communication,* vol. 34, no. 5, pp. 1086-1102, May 2016.

[9] L. Zheng, M. Ni, L. Cai, J. Pan, C. Ghosh and K. Doppler, "Performance Analysis of Group-Synchronized DCF for Dense IEEE 802.11 Networks," in *IEEE Transactions on Wireless Communications,* vol. 13, no. 11, pp. 6180-6192, Nov. 2014.

[10] D. L. Hoang, T. Hong Tran and Y. Nakashima, "Performance Evaluation of 802.11ah Physical Layer Phase Encryption for IoT Applications," *IEEE International Conference on Advanced Technologies for Communications (ATC),* Ho Chi Minh City, Vietnam, Oct. 2018.

[11] E. Bou-Harb, M. Debbabi and C. Assi, "Cyber Scanning: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials,* vol. 16, no. 3, pp. 1496-1519, Third Nov. 2013.

[12] L. Metongnon, E. C. Ezin and R. Sadre, "Efficient probing of heterogeneous IoT networks," *IFIP/IEEE Symposium on Integrated Network and Service Management (IM),* Lisbon, Portugal, May 2017.

[13] C.-W. Chang and J.-C. Chen, "UM paging: Unified M2M Paging with optimal DRX cycle", *IEEE Transaction on Mobile Computing,* vol. 16, no. 3, pp. 886-900, Mar. 2017.

[14] J. Liang, J. Chen, H. Cheng and Y. Tseng, "An Energy-Efficient Sleep Scheduling With QoS Consideration in 3GPP LTE-Advanced Networks for Internet of Things," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* vol. 3, no. 1, pp. 13-22, Mar. 2013

[15] H. S. Jang, B. C. Jung and D. K. Sung, "Dynamic Access Control With Resource Limitation for Group Paging-Based Cellular IoT Systems," in *IEEE Internet of Things Journal,* vol. 5, no. 6, pp. 5065-5075, Dec. 2018.

[16] C. H. Wei, R. G. Cheng and S. L. Tsao, "Performance Analysis of Group Paging for Machine-Type Communications in LTE Networks," in *IEEE Transactions on Vehicular Technology,* vol. 62, no. 7, pp. 3371-3382, Sept. 2013.

[17] S. Herreria-Alonso, M. Rodriguez-Perez, M. Fernandez-Veiga and C. Lopez-Garcia, "Adaptive DRX Scheme to Improve Energy Efficiency in LTE Networks With Bounded Delay," in *IEEE Journal on Selected Areas in Communications,* vol. 33, no. 12, pp. 2963-2973, Dec. 2015.

[18] K. Feng, W. Su and Y. Yu, "Design and Analysis of Traffic-Based Discontinuous Reception Operations for LTE Systems," in *IEEE Transactions on Wireless Communications,* vol. 16, no. 12, pp. 8235-8249, Dec. 2017.

[19] W. Cao, A. Dytso, G. Feng, H. V. Poor and Z. Chen, "Differentiated Service-Aware Group Paging for Massive Machine-Type Communication," in *IEEE Transactions on Communications,* vol. 66, no. 11, pp. 5444-5456, Nov. 2018.

[20] R. Harwahyu, R. Cheng and R. F. Sari, "Consecutive group paging for LTE networks supporting machine-type communications services," *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC),*, pp. 1619-1623, London, U.K, Nov. 2013.

[21] Z. Durumeric, E. Wustrow, and J. A. Halderman,"ZMap: Fast Internet-wide Scanning and Its Security Applications", *Proc. 22nd USENIX Security Symposium*, Washington, D.C., USA, Aug. 2013.

[22] R. Graham, "MASSCAN: Mass IP port scanner", Available online: https://github.com/robertdavidgraham/masscan (accessed on 16 Dec. 2020).

[23] A. Sperotto, G. Schaffrath, R. Sadre, C. Morariu, A. Pras and B. Stiller, "An Overview of IP Flow-Based Intrusion Detection," in *IEEE Communications Surveys & Tutorials,* vol. 12, no. 3, pp. 343-356, Apr. 2010.

[24] F. J. Aparicio-Navarro, K. G. Kyriakopoulos, Y. Gong, D. J. Parish and J. A. Chambers, "Using Pattern-of-Life as Contextual Information for Anomaly-

Based Intrusion Detection Systems," in *IEEE Access,* vol. 5, pp. 22177-22193, Oct. 2017.

[25] H. Hashida, Y. Kawamoto and N. Kato,"Efficient Delay-Based Internet-Wide Scanning Method for IoT Devices in Wireless LAN," in *IEEE Internet of Things Journal,* vol. 7, no. 2, pp. 1364-1374, Feb. 2020.

[26] E. Khorov et al., "Enabling the Internet of Things With Wi-Fi Halow-Performance Evaluation of the Restricted Access Window," in *IEEE Access,* vol. 7, pp. 127402-127415, Sept. 2019.

[27] F. Khan and D. Zeghlache, "Effect of cell residence time distribution on the performance of cellular mobile networks," *IEEE 47th Vehicular Technology Conference. Technology in Motion,* Phoenix, AZ, USA, 1997.

[28] 3GPP TR 37.868 V11.2.0, Study on RAN Improvements for Machine-Type Communications, Sept. 2011.

[29] S. Verma, Y. Kawamoto, H. Nishiyama, N. Kato and C. Huang, "Novel Group Paging Scheme for Improving Energy Efficiency of IoT Devices over LTE-A Pro Networks with QoS Considerations," *IEEE International Conference on Communications (ICC)*, Kansas City, MO, USA, May 2018.

[30] S. Tozlu, M. Senel, W. Mao and A. Keshavarzian, "Wi-Fi enabled sensors for Internet of things: A practical approach," *IEEE Communications Magazine,* vol. 50, no. 6, pp. 134-143, June 2012.

[31] A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi, "Internet of Things for Smart Cities," in *IEEE Internet of Things Journal,* vol. 1, no. 1, pp. 22-32, Feb. 2014.

[32] Khan, Jamil Y., Dong Chen, and Jason Brown. "A cooperative MAC protocol for a M2M heterogeneous area network." *Journal of Sensor and Actuator Networks* vol. 5, no. 3, pp. 12, Jul. 2016.

[33] C. Gomez, J. Paradells, C. Bormann and J. Crowcroft, "From 6LoWPAN to 6Lo: Expanding the Universe of IPv6-Supported Technologies for the

Internet of Things," in *IEEE Communications Magazine,* vol. 55, no. 12, pp. 148-155, Dec. 2017.

[34] T. Savolainen, J. Soininen and B. Silverajan, "IPv6 Addressing Strategies for IoT," in *IEEE Sensors Journal,* vol. 13, no. 10, pp. 3511-3519, Oct. 2013.

[35] J. Guo, C. Gu, X. Chen and F. Wei, "Model Learning and Model Checking of IPSec Implementations for Internet of Things," in *IEEE Access*, vol. 7, pp. 171322-171332, 2019

[36] P. Varadarajan and G. Crosby, "Implementing IPsec in Wireless Sensor Networks," *6th IEEE International Conference on New Technologies, Mobility and Security (NTMS),* Dubai, UAE, May. 2014.

[37] W. He and K. Nahrstedt, "An Integrated Solution to Delay and Security Support in Wireless Networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC),* Las Vegas, NV, USA, Apr. 2006.

[38] Bomin Mao, Yuichi Kawamoto, and Nei Kato, "AI-based Joint Optimization of QoS and Security for 6G Energy Harvesting Internet of Things," *IEEE Internet of Things Journal,* vol. 7, no. 8, pp. 7032-7042, Aug. 2020.

[39] Z. M. Fadlullah, C. Wei, Z. Shi and N. Kato, "GT-QoSec: A Game-Theoretic Joint Optimization of QoS and Security for Differentiated Services in Next Generation Heterogeneous Networks," in *IEEE Transactions on Wireless Communications,* vol. 16, no. 2, pp. 1037-1050, Feb. 2017.

[40] Yazar, Zeki. "A qualitative risk analysis and management tool 宴鼎RAMM." *SANS InfoSec Reading Room White Paper 11*, pp 12-32, Apr. 2002.

[41] R. I. Bonilla, J. J. Crow, L. S. Basantes and L. G. Cruz, "A Metric for Measuring IoT Devices Security Levels," *IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing,* Orlando, FL, Nov. 2017

[42] Forum of Incident Response and Security Teams (FIRST), "Common Vulnerability Scoring System version 3.1: Specification Document", Available online: https://www.first.org/cvss/specification-document (accessed on 16 Dec. 2020).

[43] National Institute of Standards and Technology, "National Vulnerability Database (NVD)", Available online: https://nvd.nist.gov/ (accessed on 16 Dec. 2020).

[44] N. Nikaein et al., "Simple Traffic Modeling Framework for Machine Type Communication," *IEEE Tenth International Symposium on Wireless Communication Systems,* Ilmenau, Germany, Aug. 2013.

[45] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," in *IEEE Journal on Selected Areas in Communications,* vol. 18, no. 3, pp. 535-547, Mar. 2000.

[46] "Draft-delcarpio-6lo-wlanah-01 - IPv6 over 802.11ah", [online] Available: https://tools.ietf.org/html/draft-delcarpio-6lo-wlanah-01.

[47] M. McQueen, W. Boyer, M. Flynn, G. Beitel, "Time-to-Compromise Model for Cyber Risk Reduction Estimation", *Proc. in Quality of Protection Workshop Springer,*, Boston, MA, U.S.A, Sept. 2005.

Acknowledgments

I am deeply indebted and grateful to my supervisor, Prof. Nei Kato, for his continuous guidance, supervision, and warm support which motivated me to carry high standard of researches. His kindness, prudence and work of ethics have made my PhD period one of the best periods of my life. I am also grateful to him for being such a patient and consistent supporter.

I would like to express my sincere gratitude to Prof. Takuo Suganuma and Prof. Hiroki Nishiyama for their great guidance which helped me to write this thesis. Their meticulous comments have been an enormous help to me.

I owe my deepest gratitude to Assoc. Prof. Yuichi Kawamoto for his continuous support and valuable advices. Without his guidance and persistent help, this thesis would not have been possible.

This thesis would not have materialized without the help of my colleagues, Tiago Koketsu Rodrigues, Yunseong Lee, and Hiroaki Hashida. I will never forget their valuable suggestions, discussions, and encouragement for the rest of my life. Also, I would like to offer my special thanks to Ms. Motoko Shiraishi for their kind support in my laboratory life. Also, I would like to express my gratitude towards past and present members of our laboratory for their continuous and friendly support over the years. Irreplaceable discussions with lab mates have helped me many times.

Furthermore, acknowledgments are also given to Japan Society of Promotion and Sciences for the research fund and allowance, which keep me focused on researches without external distraction.

Finally, I would like to express my sincere appreciation to my family who were always encouraging me to study and work hard. Without their encouragement and support, my life would have been very hard.