Network Resource Management for Delay-Sensitive Applications in Mobile Networks

August 2020

Nobuhiko Itoh

A Thesis for the Degree of Ph.D. in Engineering

Network Resource Management for Delay-Sensitive Applications in Mobile Networks

August 2020

Graduate School of Science and Technology
Keio University

Nobuhiko Itoh

Acknowledgments

I joined Associate Professor Kubo's laboratory in 2017. This dissertation is the result of three years of work at Keio University. I would like to express my gratitude to all those who gave me the opportunity to accomplish this dissertation.

First of all, I would like to express the deepest appreciation to my supervisor Associate Professor Dr. Ryogo Kubo. He was kind enough to discuss with me at any time even though he is busy. His constructive advices and his significant supports throughout this work are a great help to me.

I am deeply grateful to Professor Dr. Masaaki Ikehara, Professor Dr. Yukitoshi Sanada, and Professor Dr. Hiroaki Nishi, for giving me a lot of instructive advices and helpful discussions. Their supports have been of great value in this study.

I owe my warm gratitude to Mr. Takanori Iwai and Dr. Kozo Satoda, the System Platform Research Laboratories, NEC Corporation, and Mr. Motoki Morita, the Smart Infrastructure Division, NEC Corporation, for giving me a lot of helpful advices. I wish to thank Mr. Dai Kanetomo and Dr. Eiji Takahashi, the System Platform Research Laboratories, NEC Corporation. They gladly consented to study at Keio University.

I owe my sincere gratitude to Professor Dr. Takahiro Matsuda, Tokyo Metropolitan University, and Professor Dr. Takefumi Hiraguri, Nippon Institute of Technology, for giving me valuable comments through technical meetings.

During this study, I have collaborated with many colleagues including those who

could not be mentioned here. Especially, I am grateful to all working with me at the System Platform Research Laboratories, NEC Corporation, for giving me the chance to study at Keio University.

I am warmly grateful to my family who gave me time to complete this dissertation. My wife gave me a lot of warm words even when I was working until late in the night. My children encouraged me with her or his smile. I would like to express my gratitude again to all those who supported my research.

August 2020

Nobuhiko Itoh

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

This section describes the history of mobile networks and delay-sensitive applications.

1.1.1 Mobile Networks

Mobile networks have continuously evolved over the decades. The main radio technologies and signature services of each generation are listed in Table 1.1. The initial mobile networks, called first generation (1G) networks, are realized by analog communication. The main service of 1G is voice. In second generation (2G) networks and those that followed, digital communication is used instead of analog communication. This digital communication led to the concept of packet-based communication, and e-mail service was launched. Third generation (3G) networks are based on an all Internet protocol (IP) network infrastructure and use advanced radio access technologies, such as multiple-input and multiple-output (MIMO). With the evolution of the mobile network, in fourth generation (4G) networks, represented by long term evolution (LTE) and LTE-Advanced, video streaming services became available on mobile phones.

The history of mobile networks in detail is shown in Fig. 1.1, particularly in terms

Table 1.1: Main technologies and signature services of each generation.

Generation	Main technologies	Signature services
First generation (1G) 1980 - 1990 years	Analog	Voice
Second generation (2G) 1990 - 2000 years	Digital	Packet communication (e-mail)
Third generation (3G) 2000 - 2010 years	Digital	IP communication
Fourth generation (4G) 2010 - 2020 years	LTE and LTE-Advanced	High bit rate services such as video streaming

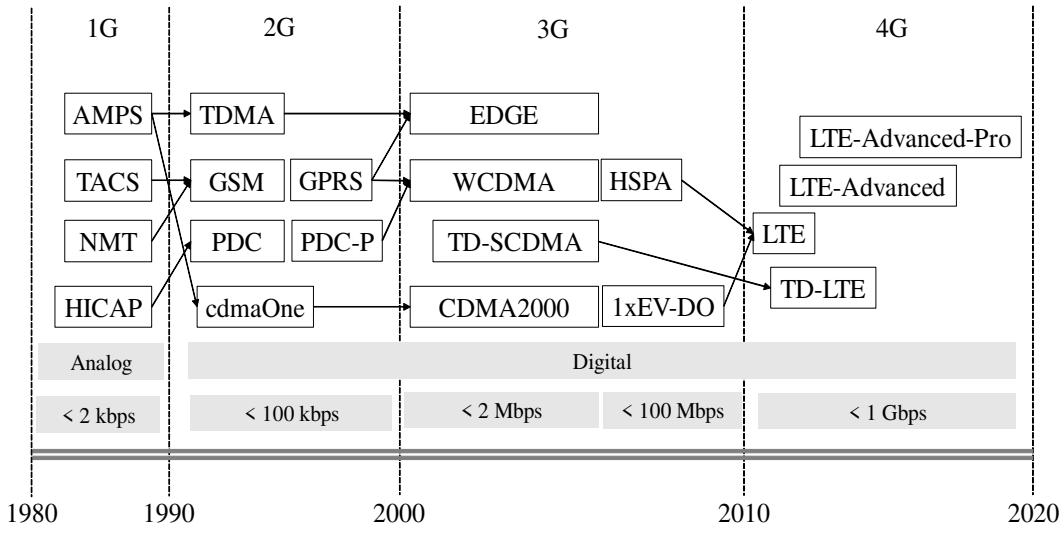

Figure 1.1: History of mobile networks.

of radio access technologies.

1G was applied to commercial automobile telephones around 1980 [2,3]. In Japan, Nippon Telegraph and Telephone Corporation (NTT) started an automobile telephone service with a high capacity (HICAP) method in 1979. Europe also started the commercialization of mobile networks with the Nordic mobile telephony (NMT) 450 method in 1981 [4]. In the United States, the commercialization of mobile networks

with advanced mobile phone service (AMPS) method was started in 1983 [5]. The United Kingdom started the commercialization of mobile networks with the total access communication system (TACS), which is an AMPS-based technology. 1G was based on analog transmission and was limited to voice services. Since the mid-1980s, the downsizing of mobile terminals has progressed, enabling us to carry the mobile terminal.

2G, which emerged in the early 1990s, introduced digital transmission via radio link. The main service of 2G was still voice. In addition, 2G provided users with data services. The representative technology of 2G is the global system for mobile communications (GSM) [6], jointly developed by a large number of European countries. In the United States, digital AMPS (D-AMPS) [7] was developed and introduced in 1991. In 1993, personal digital cellular (PDC) was introduced in Japan. These methods can both be categorized as time division multiple access (TDMA), which allows multiple users to utilize the same frequency channel by dividing the signal into different time slots [8]. Code-division multiple access (CDMA) was developed in 1996 [9]. CDMA is an example of multiple access, where multiple transmitters can send data simultaneously over a single radio channel. This characteristic of CDMA allows multiple users to share a band of frequencies. As the downsizing of the mobile terminal has been further progressed, users have been able to carry smaller and lighter mobile terminals. Digitalization has allowed users to receive not only voice services but also data services. The circuit-switched system was used to realize data services; further, the data rate of the system ranges from 9.6 kbps to tens of kbps. Packet-switching systems then came to use in the main system: PDC packet (PDC-P) [10] and general packet radio service (GPRS) [11].

3G was introduced in the early 2000s. 3G provides users with high-quality

mobile services and high-speed transmission. 3G uses the international mobile telecommunications-2000 (IMT-2000) [12] specifications by the International Telecommunication Union (ITU) as radio access technology. The representative technologies of IMT-2000 are wideband CDMA (WCDMA) [13] and CDMA2000 [14]. The maximum transmission rates of WCDMA and CDMA2000 are 384 kbps and 144 kbps, respectively. GSM evolution (EDGE), which is GSM-based radio technology, enhanced the data rate to a maximum of 384 kbps. Time-division synchronous CDMA (TD-SCDMA) is an air interface found in China as an alternative to W-CDMA. Users can send text, pictures, and movies with this technology in 3G. The concept of the best effort that adaptively selects transmission rate in accordance with radio quality was introduced as 3.5G. In 3.5G, CDMA2000 was enhanced, and the enhanced CDMA2000 was called evolution data optimized (1xEV-DO) [15]. 1xEV-DO was introduced in 2003 in Japan. The maximum downlink rate of 1xEV-DO is approximately 2.4 Mbps. WCDMA was also enhanced, namely as high-speed packet access (HSPA) [16]. HSPA was introduced in 2006 in Japan. The maximum downlink rate of HSPA is approximately 14 Mbps.

In 4G, which is represented by LTE and LTE-Advanced, LTE has followed in the steps of HSPA, providing users with higher efficiency and further enhanced user experience [17, 18]. For example, its targets for maximum throughput in terms of uplink and downlink are 50 Mbps and 100 Mbps, respectively, when bandwidth is 20 MHz. The target delay budget of the radio section is 10 ms. In LTE release 8, orthogonal frequency-division multiplexing (OFDM) is the downlink multiple access schemes, while single-carrier frequency-division multiple access (SC-FDMA) is the uplink multiple access scheme. Bandwidth extension in LTE-Advanced is supported by carrier aggregation. Carrier aggregation allows mobile networks to provide band-

widths of up to 100 MHz, enabling a maximum transmission rate in excess of 1 Gbps in downlink and 500 Mbps in uplink [19]. LTE-Advanced-Pro is a development of LTE-Advanced that supports transmission rates in excess of 3 Gbps using 32-carrier aggregation. Time-division LTE (TD-LTE) that is familiar with TD-SCDMA does not require paired spectrum because both transmit and receive occur on the same channel. It is possible to dynamically change uplink and downlink radio capacity to match demand.

The developments of mobile network components are described in Fig. 1.2. The 1G network consists of a radio access network (RAN), circuit switching core network (CN), and fixed telephone network. The RAN includes a base station to connect mobile phones. Mobile phones connect to the fixed telephone network via RAN and the circuit switching CN. The 2G network components are the same as the 3G network components. The key change in the system from 1G is that a packet switching domain is added in the 2G CN. The packet switching domain connects to the Internet. Therefore, mobile phone users can receive some application services provided by content providers through the packet switching CN. 4G networks based on LTE support only the packet switching domain. In 4G, telephone services with the circuit switching domain were discontinued, and voice over LTE (VoLTE) services that send voice with the packet switching domain have been started.

Discussions on fifth generation (5G) mobile networks began during 2012. In 5G, new radio (NR) access technology has been discussed [20–22]. 5G NR was designed to be the global standard for the air interface of 5G networks. The Third Generation Partnership Project (3GPP) provides the technical details of NR in the 3GPP specification 38 series [23]. In 5G NR, it is assumed that two different frequency ranges are utilized. First, there is frequency range one (FR1), which includes sub-6

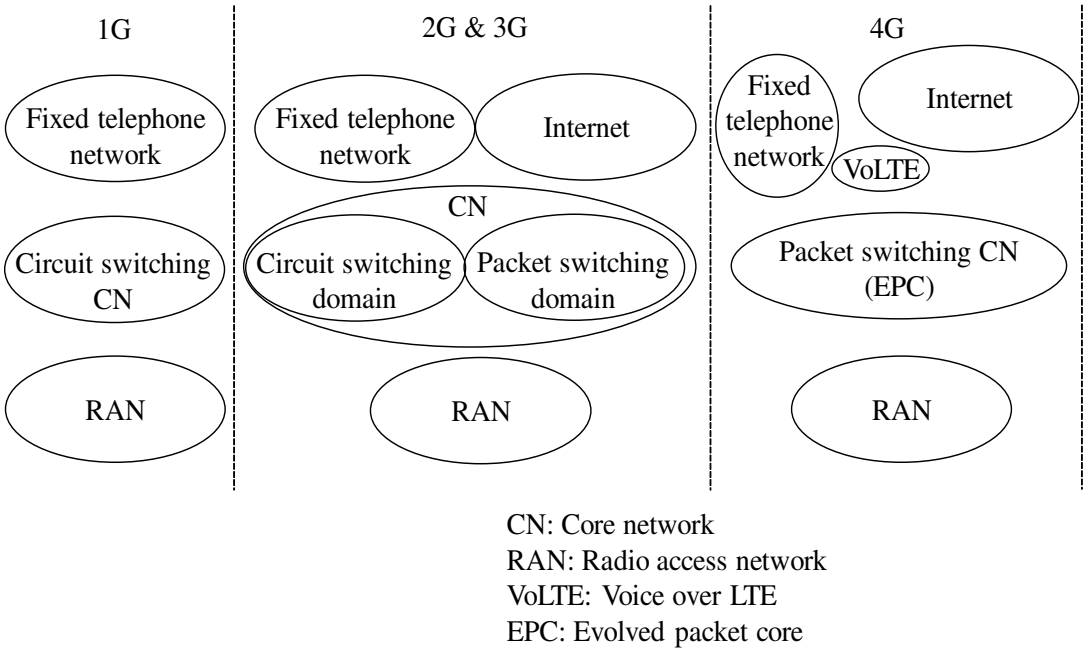

Figure 1.2: Developments of mobile network components.

GHz frequency bands from 410 Mhz to 7125 MHz [24]. The other is frequency range two (FR2), which includes frequency bands from 24.25 GHz to 52.6 GHz, called millimeter wave frequencies [25]. Although high-frequency bands, such as this millimeter wave range, have shorter ranges, FR2 provides higher available bandwidths than bands in FR1.

5G NR has two deployment modes: non-standalone (NSA) mode and standalone (SA) mode, as shown in Fig. 1.3. The main difference between the NSA and SA modes is that NSA mode anchors the control signaling of 5G for a user equipment (UE) to the 4G base station, and in SA mode, the 5G base station is directly connected to the 5G core network. In SA mode, both the control signaling and data pass through the 5G base station to access the 5G core network. The initial phase of 5G NR is assumed to be constructed with the NSA mode. 5G NR is then constructed with the SA mode. The SA mode includes the new 5G packet core architecture instead of relying on the

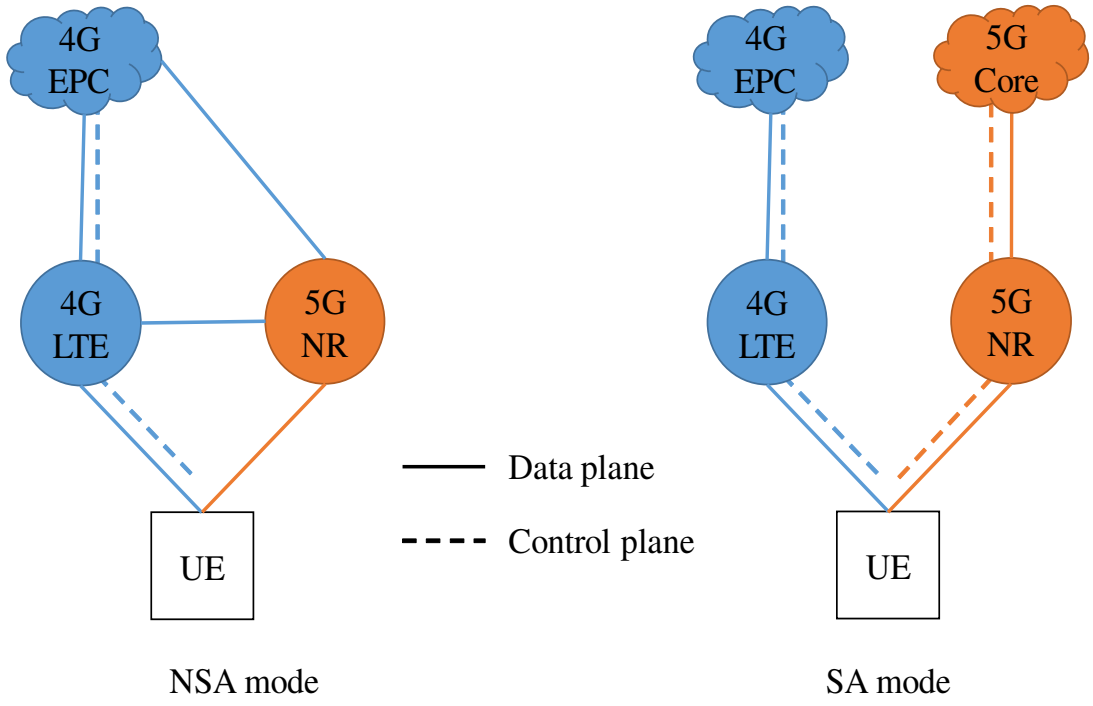

Figure 1.3: Deployment mode of 5G NR (NSA and SA modes).

evolved packet core (EPC). When the migration has completed, 5G NR can realize better efficiency and lower cost and advance the development of new use cases.

In addition, 5G use cases also have been discussed. The International Telecommunication Union Radiocommunication Sector (ITU-R) has defined three main use cases for 5G, as shown in Fig. 1.4; they are enhanced mobile broadband (eMBB), ultra-reliable low latency communications (URLLC), and massive machine-type communications (mMTC) [26].

eMBB provides greater bandwidth complemented with moderate latency improvements on 5G NR. This will help to develop mobile broadband use cases, such as emerging augmented reality (AR), virtual reality (VR), ultra-high definition (UHD), and 360-degree streaming video. URLLC focuses on delay-sensitive applications, such as self-driving cars and remote management. Industries have high expectations for URLLC. mMTC focuses on services that have strong requirements for connection

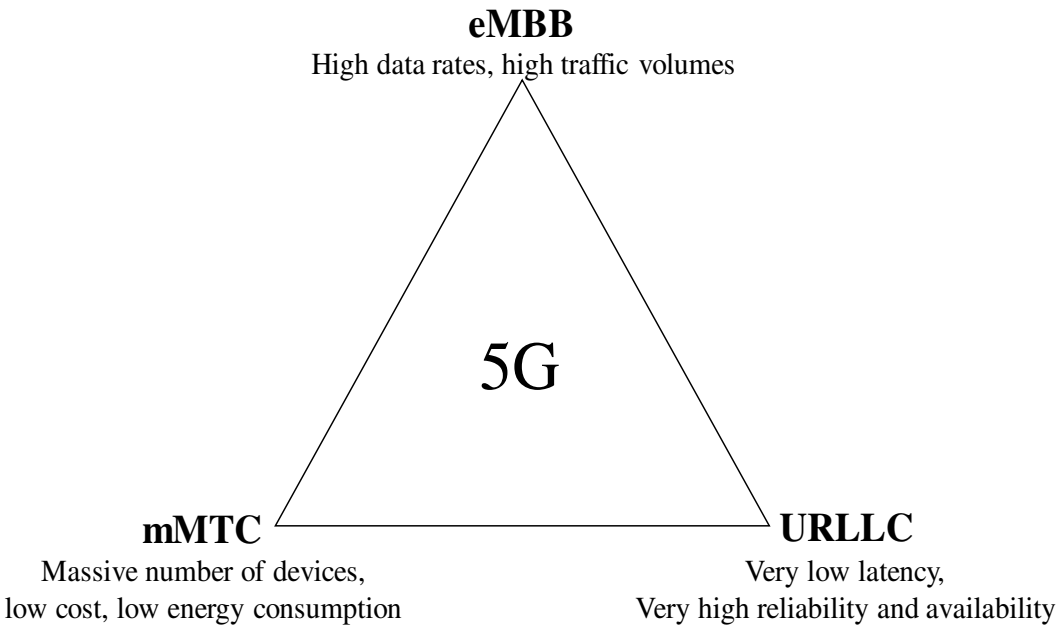

Figure 1.4: High-level 5G use case classification.

Table 1.2: Orders of expected response times.

Case	Reaction time
Web browsing	1 s
Voice	100 ms
TV	10 ms
VR	1 ms

density, such as smart cities and Internet of Things (IoT).

1.1.2 Delay-Sensitive Applications

Mobile networks enable devices to use applications for remote control in addition to conventional applications, such as voice and video streaming. The order of expected response time for each application is shown in Fig. 1.2 [27].

According to [28], the ITU Telecommunication Standardization Sector (ITU-T)

argues the following. In the case of interactive web browsing, the maximum tolerable response time is approximately 1 s. When response time for interactive web browsing is less than a few hundred milliseconds, users find good satisfaction in the interactive web browsing.

The maximum tolerable response time for voice is approximately 100 ms. As an example of VoLTE, to realize natural conversation, voice packets must be delivered within 100 ms. If the interval between pictures when watching television (TV) is less than 10 ms, the user will be satisfied with the quality of the TV. It is expected that the latency of video streaming is also less than 10 ms.

In the case of a moving head-mounted virtual reality (VR) device, users are not satisfied by the quality of service when the time until additional information is displayed on the VR device is greater than human reaction time; the required response time is within 1 ms when mobile networks are used to acquire additional information displayed on a VR device.

Mobile networks enable individual devices, such as robots, vehicles, drones, and sensors, to exchange real-time information to control devices remotely. As an example of remote control, there is automation in industries. A 3GPP report identified a set of performance targets for industrial automation using 5G [29]. Table 1.3 lists the industrial automation performance requirements for 5G [1]. As shown in Fig. 1.3, all use cases require high availability and low latency. Service availability, described as Availability in Fig. 1.3, refers to the percentage of time for the end-to-end communication service is delivered. Cycle time is the time from when a device generates a command to when the device receives an acknowledgment message from a sensor or actuator. Therefore, it is expected that the latency over the network for transmission must be at least lower than the cycle time.

Table 1.3: Performance targets for industrial automation [1].

Use case		Availability	Cycle time	Assumed payload size	# of devices	Assumed service area
Motion control	Printing machine	> 99.9999 %	< 2 ms	20 Byte	> 100	100 m * 100 m * 30 m
	Machine tool	> 99.9999 %	< 0.5 ms	50 Byte	< 20	15 m * 15 m * 3 m
	Packaging machine	> 99.9999 %	< 1 ms	40 Byte	< 50	10 m * 5 m * 3 m
Mobile robots	Cooperative motion control	> 99.9999 %	1 ms	40 -250 Byte	100	< 1 km^2
	Video-operated remote control	> 99.9999 %	10 - 100 ms	15 - 150 Byte	100	< 1 km^2
Mobile control panels with safety functions	Assembly robots or milling machines	> 99.9999 %	4 - 8 ms	40 - 250 Byte	4	10 m * 10 m
	Mobile cranes	> 99.9999 %	12 ms	40 - 250 Byte	2	40 m * 60 m
Process automation		> 99.99 %	> 50 ms	Varies	10000 devices per km^2	

When controlling industrial robots moving rapidly, an end-to-end delay of below 1 ms per sensor is required because the cycle time must be below 2 ms [27]. Figure 1.5 gives a breakdown of cycle time. Assuming that each delay of robotic device processing and cloud and application processing is 250 μs, the robotic motion control system equates to 500 μs of one-way air interface delay. Using mini-slots, which are the minimum scheduling unit [30,31], 5G NR can potentially deliver a packet in below 500 μs.

As another example of remote control, there is a connected car application for avoiding road traffic collision. Road traffic collisions are an extremely serious issue worldwide. According to the World Health Organization (WHO), the number of road

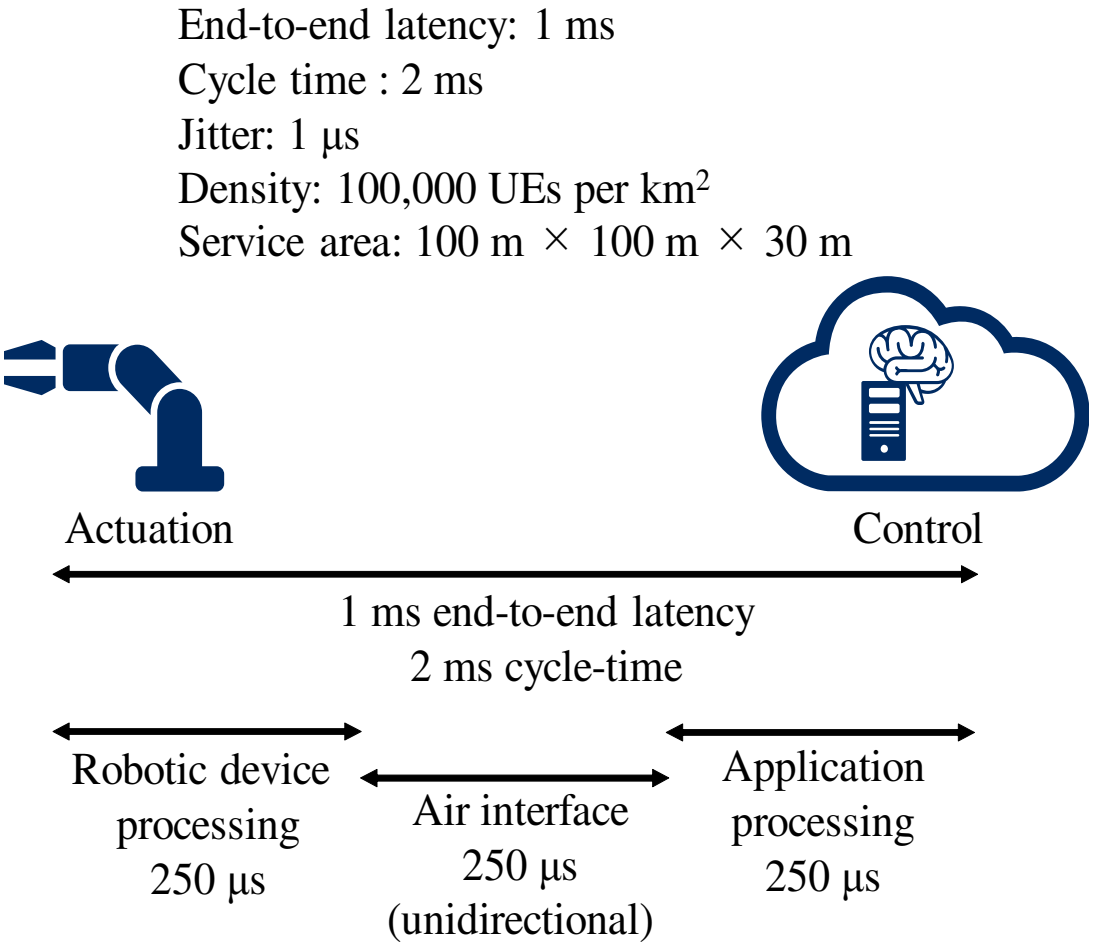

Figure 1.5: Delay breakdown for robotic motion control.

traffic deaths has now grown to 1.25 million per year [32]. In Japan, where drivers keep to the left side of the road, the number of fatalities is 4,000 per year, and road traffic collisions occurring in blind spots, such as right-turn collisions and rear-end collisions, account for 74% of all accidents [33].

In an attempt to decrease the occurrence of road traffic collisions, driving safety systems with sensors, such as vehicle-mounted cameras and vehicle radar, for detecting pedestrians and vehicles have been deployed. However, these systems cannot detect vehicles and pedestrians in blind spots, such as at sharp bends and blind intersections.

To mitigate road traffic collisions at intersections, connected cameras are allocated

around an intersection. Mobile networks, such as LTE networks, have attracted a great deal of attention as platforms for detecting vehicles and pedestrians at an intersection [34–36].

By using mobile networks, a connected car service can collect real-time information, such as an image from around the intersection and the locations of vehicles, from connected devices, such as road side cameras and vehicles. This real-time information can help drivers to avoid road traffic collisions. For example, a road traffic collision avoidance scheme with road-side cameras was proposed in [37]. In London [38] and Beijing [39], many cameras have been deployed for surveillance in public spaces. In the future, these cameras may be utilized for vehicle traffic control.

For guaranteeing effective connected car services, it is important to deliver a data block within a certain maximum tolerable delay (called a deadline in this dissertation). The network requirements for connected cars are listed in Table 1.4. This research assumes LTE vehicle-to-everything (V2X) services, such as vehicle-to-vehicle (V2V), vehicle-to-pedestrian (V2P), and vehicle-to-infrastructure (V2I). 3GPP stipulates that the deadline for an urban intersection is 100 ms and that the arrival ratio within the deadline is more than 95% [40].

1.2 Orientation of Research

This research aims to establish network resource management of mobile networks for delay-sensitive applications. It is assumed that IoT devices using delay-sensitive applications, such as drones, vehicles, robots, and monitoring cameras, send a request message to a server for delay-sensitive applications and receive a reply message from the server, as shown in Fig. 1.6. The request message includes information such as

Table 1.4: Network requirements for a connected car.

	Effective distance	Absolute speed of a UE supporting V2X Services	Relative speed between 2 UEs supporting V2X Services	Maximum tolerable latency	Minimum radio layer message reception reliability (probability that the recipient gets it within 100ms) at effective distance	Example Cumulative transmission reliability
#1 Suburban and major road	200 m	50 km/h	100 km/h	100 ms	90%	99%
#2 Freeway and motorway	320 m	160 km/h	280 km/h	100 ms	80%	96%
#3 Autobahn	320 m	280 km/h	280 km/h	100 ms	80%	96%
#4 (NLOS and urban)	150 m	50 km/h	100 km/h	100 ms	90%	99%
#5 Urban intersection	50 m	50 km/h	100 km/h	100 ms	95%	-
#6 campus and shopping area	50 m	30 km/h	30 km/h	100 ms	90%	99%
#7 Imminent crash	20 m	80 km/h	160 km/h	20 ms	95%	-

location information and sensing information. The reply message includes control information for controlling the behavior of the IoT devices. In this research, a flow that consists of both request flow and reply flow is defined as cycle flow. This research contributes to guaranteeing deadline constraints of the cycle flow. The main research topics discussed in this dissertation are as follows:

- Scheduler technique on a base station
- Bandwidth assignment technique on a base station and gateway

The scheduler technique for delay sensitive applications on a base station includes a priority control and deadline coordination function. In this research, priority control means a method for deciding the priority of the cycle flow in accordance with the application characteristics, such as deadline and data size, and the network conditions,

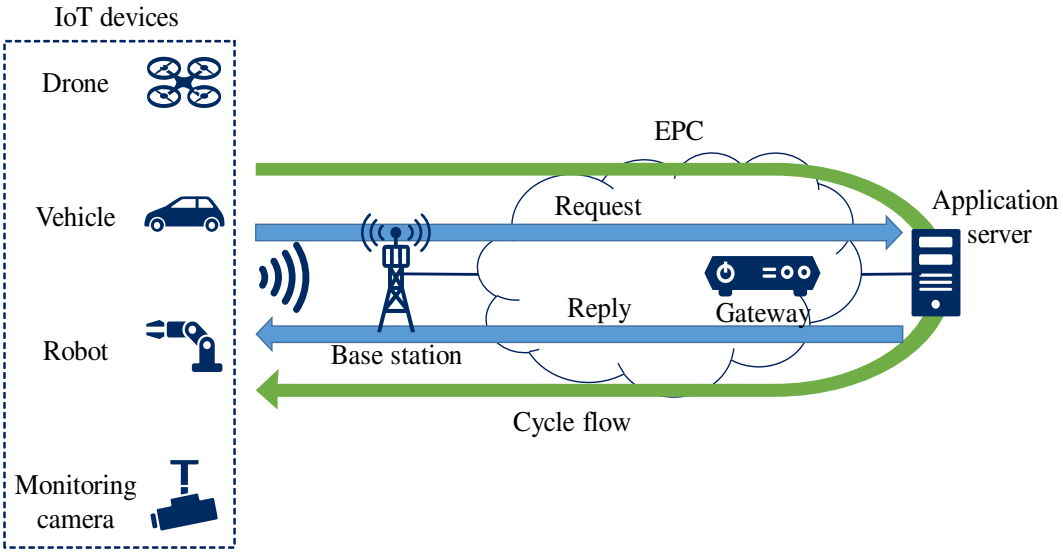

Figure 1.6: Cycle flow.

such as the fluctuation of radio quality. The deadline coordination function is a method for allocating uplink and downlink deadlines to each link in accordance with the deadline of the cycle flow and congestion levels of uplink and downlink radio sections.

The bandwidth assignment technique on a base station and gateway assigns a bandwidth with maximum bit rate (MBR) control based on throughput feedback, which is intended for uplink flow to collect information from various IoT devices.

1.3 Chapter Organization

The chapter organization is illustrated in Fig. 1.7. The following chapter describes the mobile network architecture assumed in this research and conventional network resource management methods, i.e., scheduling techniques with priority control on evolved NodeB (eNB) and bandwidth assignment techniques with MBR control on eNB and the serving gateway and packet data network gateway (S/P-GW).

Chapter 3 proposes a scheduler implemented on eNB for cycle flow. The proposed scheduler decides the priority in accordance with the application characteristics, such as deadline and data size, and the network conditions, such as the fluctuation of radio quality. Network simulations confirm that the proposed scheduler achieves better performance than conventional schedulers. The scheduler proposed in Chapter 3 was first presented in [41].

Chapter 4 proposes a method for allocating uplink and downlink deadlines to each link in accordance with the deadline of the cycle flow and congestion levels of uplink and downlink radio sections. Network simulations confirm that the introduction of the proposed method into the scheduler proposed in Chapter 3 achieves better performance than the proposed scheduler alone. The method proposed in Chapter 4 was first presented in [42].

Chapter 5 proposes a bandwidth assignment method with MBR control based on throughput feedback, which is intended for uplink flow to collect information from various IoT devices. The proposed bandwidth assignment method enables cellular networks to assign adaptive bandwidth to uplink flow on a base station and gateway. It is clarified that a conventional model-based MBR control method achieves worse performance in the presence of disturbances, such as radio quality fluctuation. Numerical simulations confirm that the proposed bandwidth assignment method achieves better performance because the feedback controller maintains the effective throughput at a target level by considering disturbances. The bandwidth assignment method proposed in Chapter 5 was first presented in [43].

Finally, Chapter 6 summarizes the findings of this research.

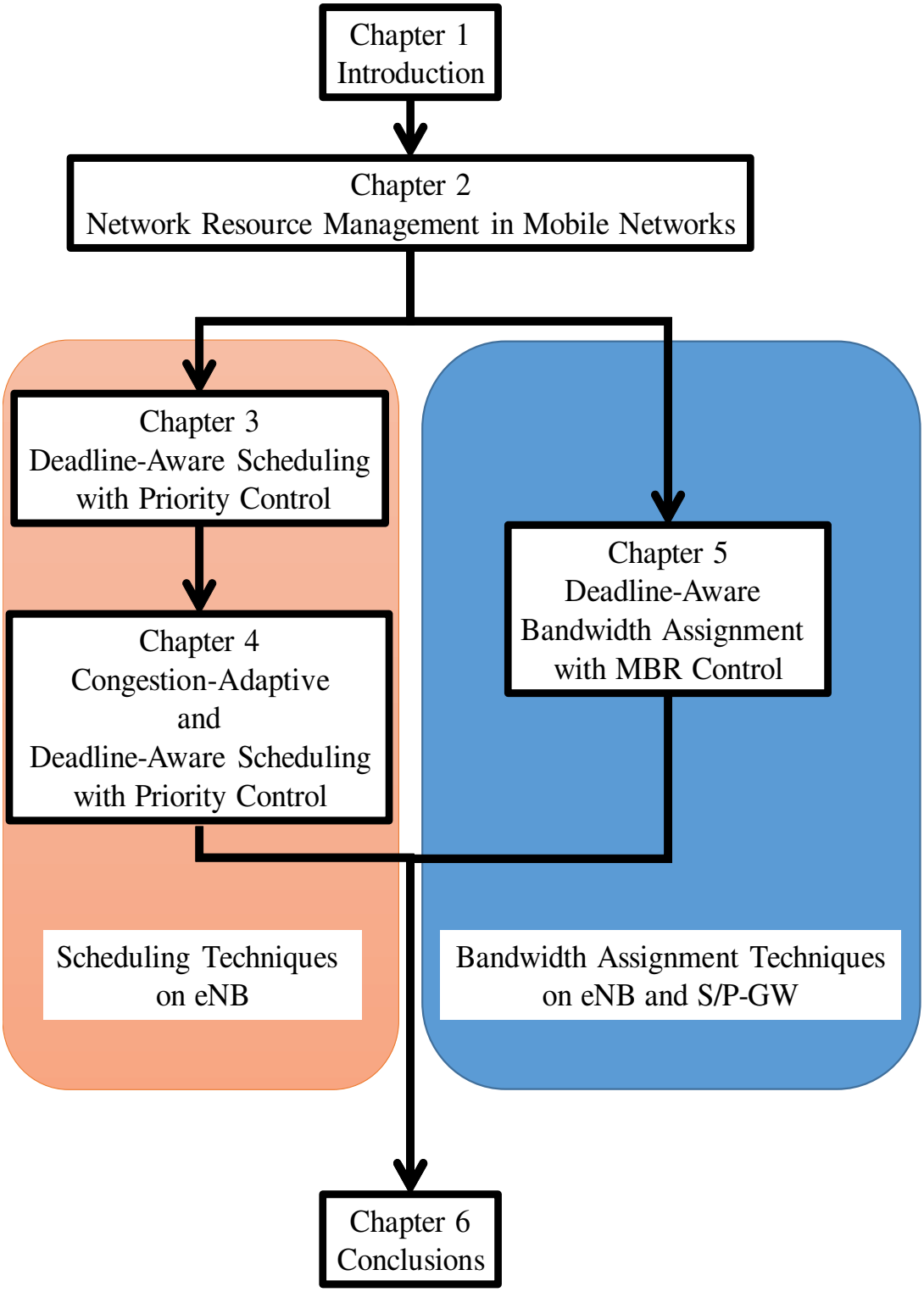

Figure 1.7: Chapter organization.

Chapter 2

Network Resource Management in Mobile Networks

2.1 Architecture

The LTE reference architecture [44], called an evolved packet system (EPS), is shown in Fig 2.1. The LTE reference architecture consists of UE, an evolved universal terrestrial radio access network (E-UTRAN) [18], which deals with technology related to a radio access network, and an evolved packet core (EPC), which deals with technology related to a core network. As shown in Fig. 2.1, EPC entities are serving gateway (S-GW), packet data network gateway (P-GW), mobility management entity (MME), home subscriber server (HSS), policy and charging rules function (PCRF), subscriber profile repository (SPR) [45], offline charging system (OFCS) [46], and online charging system (OCS) [47]. A packet data network (PDN) is an internal or external Internet protocol (IP) domain of the operator that a UE wants to communicate with, and it provides a UE with services, such as the IP multimedia subsystem (IMS).

A UE connects to an eNB over the LTE-Uu interface and transmits and receives data from an application server on a PDN. The eNB provides UEs with radio interfaces and

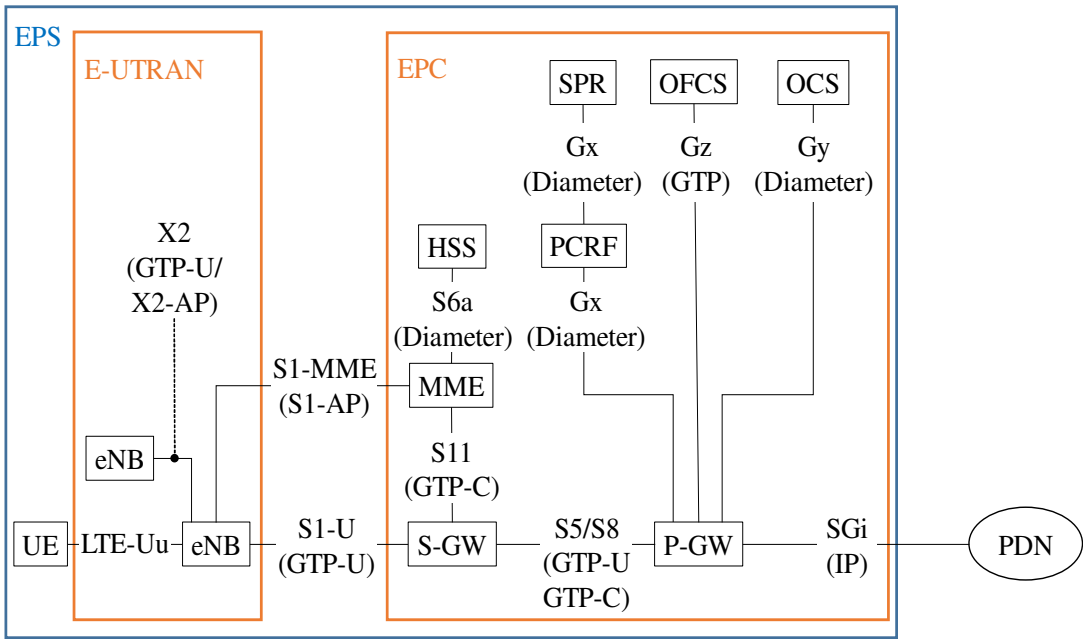

Figure 2.1: LTE reference architecture.

performs radio resource management functions, including power control, scheduling, handover, connection monitoring, traffic measurement, and connection establishment.

The roles of the EPC entities are listed in Table 2.1.

Table 2.1: EPC entities.

Entity	Description
S-GW	An S-GW terminates the interface towards the E-UTRAN. It serves as the local mobility anchor point of data connections for inter-eNB and inter-3GPP handover.
P-GW	A P-GW serves as a gateway to access to an external PDN. It assigns an IP address from the address space of the PDN to a UE. It serves as the mobility anchor point for handover between 3GPP and non-3GPP. It performs charging based on the policy and charging control (PCC) rules provided by a PCRF. Mobile operator can co-locate an S-GW and P-GW into one device, which is called an S/P-GW.
MME	An MME is the main control entity for the E-UTRAN. The MME communicates with an HSS for user authentication and user profile download. The MME provides UEs with EPS mobility management, including paging and handover, and EPS session management functions.
HSS	An HSS is a central database where user profiles are stored. The HSS provides the MME with user authentication information and user profiles.
PCRF	The role of a PCRF is policy and charging control. The PCRF makes policy decisions based on subscription information provided by SPR and provides the P-GW with the PCC rules, such as quality of service (QoS) and charging rules.
SPR	An SPR provides the PCRF with subscription information, such as access profile per subscriber.
OFCS	An OFCS provides charging information based on charging data record (CDR), which is a formatted collection of information about a chargeable telecommunication event. In general, offline charging is used in postpaid plans.
OCS	An OCS is a system for charging based on service usage in real-time. In general, online charging is used in prepaid plans.

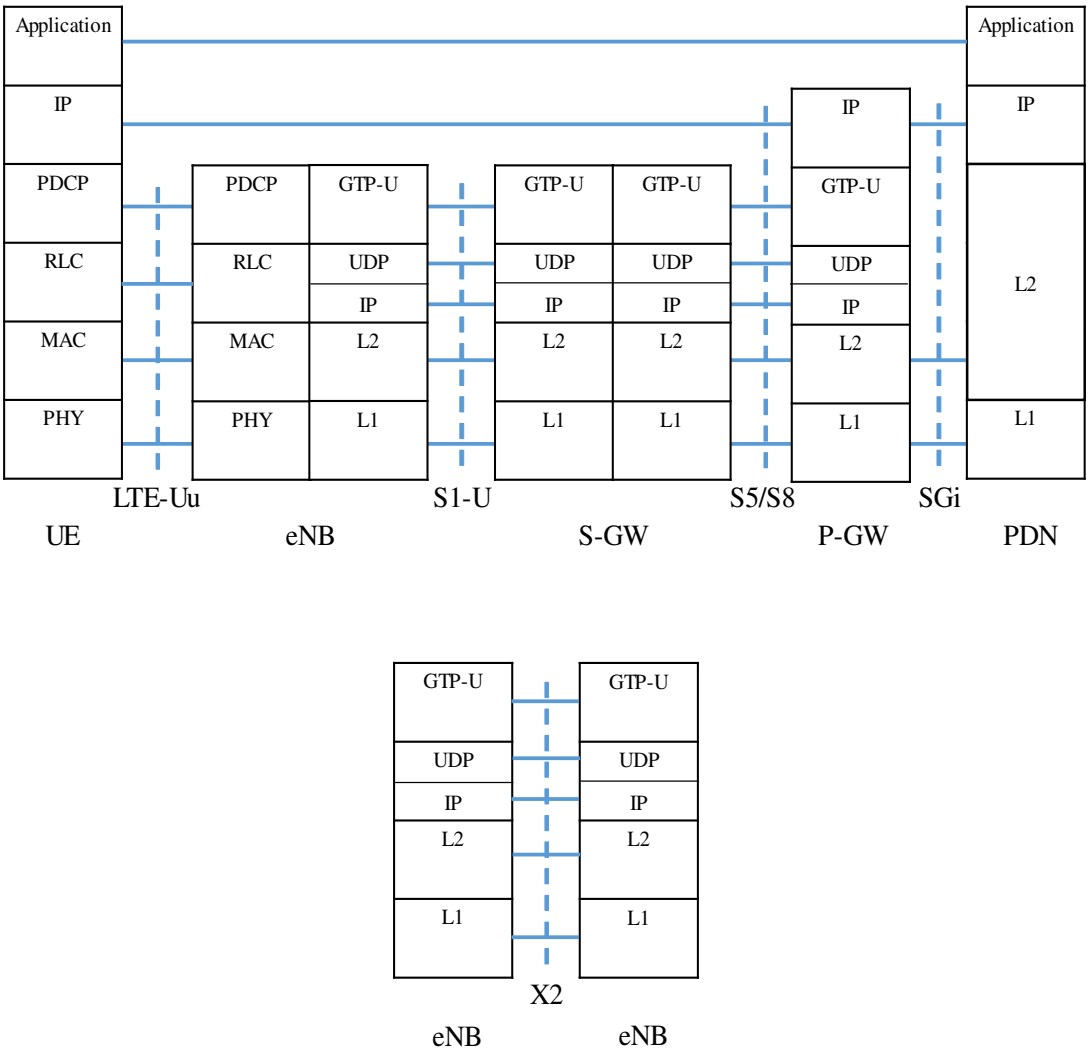

Figure 2.2: Protocol stack for user plane.

Figures 2.2 and 2.3 show protocol stacks for user plane and control plane, respectively. L1 and L2 shown in Figs. 2.2 and 2.3 represent layer 1 and layer 2, respectively.

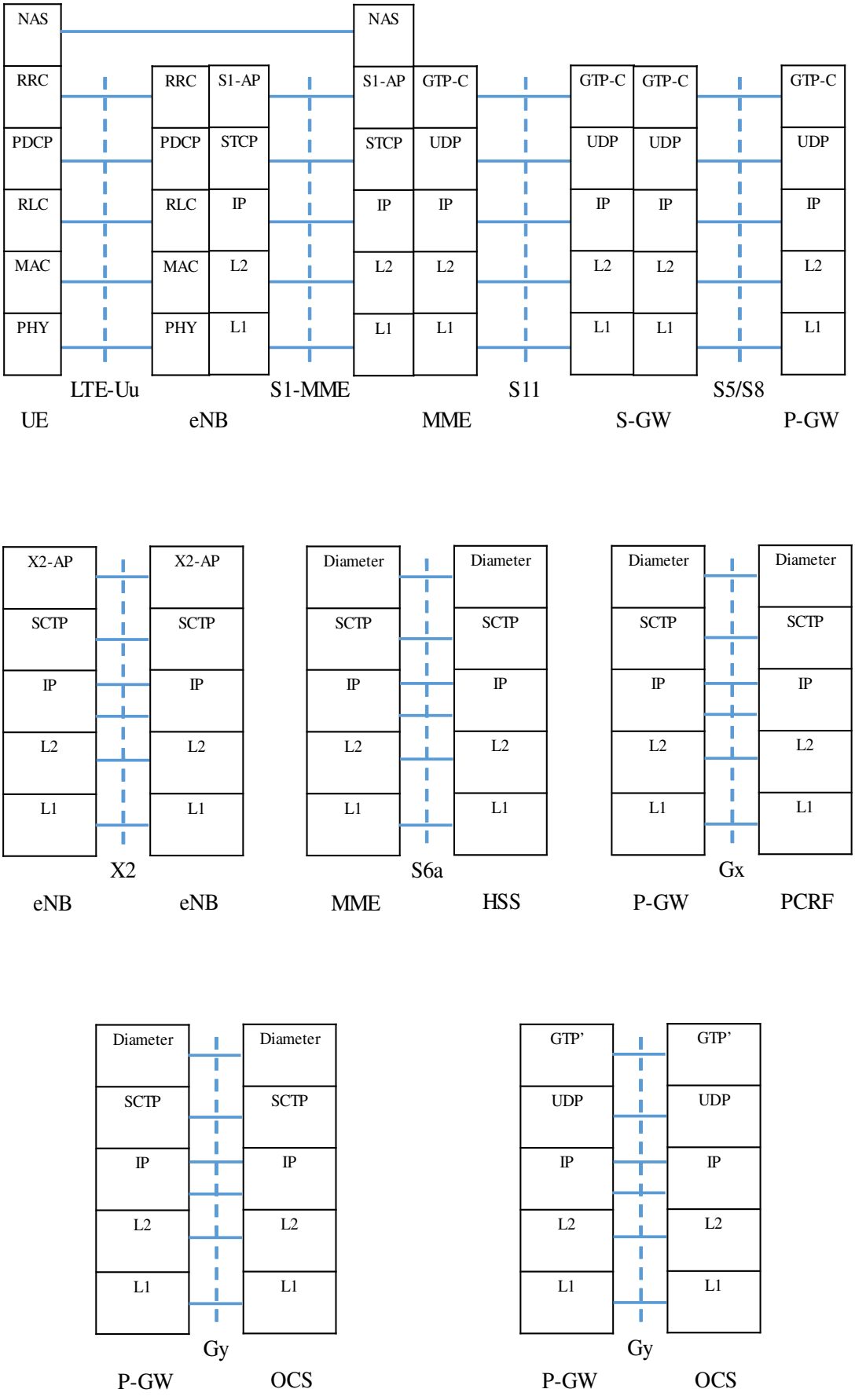

Figure 2.3: Protocol stack for control plane.

The LTE-Uu interface is an interface for the user and control planes between a UE and an eNB. The connection for the control plane over LTE-Uu is the radio resource control (RRC) connections represented by signaling radio bearers (SRBs). In addition, a non-access stratum (NAS) protocol over the LTE-Uu interface performs mobility management and bearer management functions. The data plane connection over LTE-Uu is the logical channels represented by data radio bearers (DRBs). The packet data convergence protocol (PDCP) over the LTE-Uu interface performs header compression, packet reordering, and retransmission to provide efficient data transmission over the radio section. In the transmission, the radio link control (RLC) protocol over the LTE-Uu interface constructs an RLC protocol data unit (PDU) and provides the medium access control (MAC) layer with the RLC PDU. The RLC protocol on the receiving side constructs the PDCP PDU from the received RLC PDU. The RLC protocol has three modes for reliability: transparent, acknowledged, and unacknowledged.

The MAC layer is located between the RLC layer and physical (PHY) layer. The MAC protocol supports multiplexing and de-multiplexing between logical channels and transport channels, as the MAC protocol is connected to the RLC layer through logical channels and to the PHY layer through transport channels. The MAC layer supports QoS control by scheduling with priority. The MAC scheduler dynamically allocates radio resources to UE in accordance with each priority.

The X2 interface is an interface for the user and control planes between two eNBs. The X2 interface is mainly used for X2 handover between two eNBs. When X2 handover is executed, X2-AP over the X2 interface is used as a protocol for the control plane. X2-AP enables an LTE network to support UE mobility. Then, GPRS tunneling protocol user plane (GTP-U) over the X2 interface is used as a protocol for the data

plane to forward data from source eNB to target eNB.

The S1-U interface is an interface for the user plane between an eNB and S-GW. A GTP-U protocol over the S1-U interface is used to transmit user data. The S1-MME interface is an interface for the control plane between an eNB and MME. The S1-AP protocol over the S1-MME is used to control the eNB. The S1-AP protocol supports UE context management. The S1-AP protocol delivers the initial UE context to the eNB to set up an E-UTRAN radio access bearer (E-RAB). Modification or release of the UE context is also managed by the S1-AP protocol.

The S11 interface is an interface for the control plane between an S-GW and MME. A GPRS tunneling protocol control plane (GTP-C) over the S11 interface is used to exchange control messages. The GTP-C protocol is a protocol for supporting the exchange of control information for creation, modification, and termination of GTP tunnels. The GTP-C protocol creates tunnels for data forwarding in the case of handover. The GTP-C protocol is used over the S5 and S8 interfaces for the same purposes.

S5 and S8 are the same interface, and both interfaces are interfaces for the control plane and user plane between an S-GW and P-GW. The S5 interface is used in the network internally, while S8 is used when roaming between different operators. A GTP-U protocol over the S5 or S8 interface is used to forward user data, and a GTP-C protocol is used for the control plane.

A diameter protocol over S6a, Sp, Gx, Gy, and Gz is used for the control plane. The diameter protocol is a protocol for authentication and authorization. S6a in an interface for the control plane between an HSS and MME. The S6a interface is used to exchange user subscription and authentication information. The Sp interface is an interface for the control plane between a PCRF and SPR. The Gx interface is an

interface for the control plane between a P-GW and PCRF. The PCRF sends policy control and charging rules to the P-GW over the Gx interface to support both QoS and charging controls. The Gy interface is an interface for the control plane between an OCS and P-GW. The Gy interface is used to transfer charging-related information.

The Gz interface is an interface for the control plane between an OFCS and P-GW. The Gz interface enables the transport of service data flow-based offline charging information, and then the GTP' protocol is used over the Gz interface. The GTP' protocol supports CDR transfer from the P-GW to the OFCS.

The SGi interface is defined between a P-GW and PDN for the control and user planes. An IP packet without the GTP header in the user plane is sent to the PDN. For the control plane, a dynamic host configuration protocol (DHCP) and remote authentication dial-in user service (RADIUS) over the SGi interface are used.

2.2 Scheduling Techniques on an eNB

In LTE networks, a scheduler is implemented at the MAC layer of an eNB, as shown in Fig. 2.4. The MAC scheduler consists of functions for priority control and resource allocation. The priority control function decides the priority of each UE in accordance with a QoS policy from the PCRF, the radio quality from the UE, and the amount of data waiting for transmission, as described in the buffer status report. The resource allocation function allocates radio resources to a UE in accordance with each priority so that it can transmit and receive data.

Conventional MAC schedulers have two types: deadline-unaware and deadline-aware. Deadline-unaware schedulers have been proposed to improve metrics such as fairness and system throughput. A proportional fair (PF) scheduler [48, 49] improves

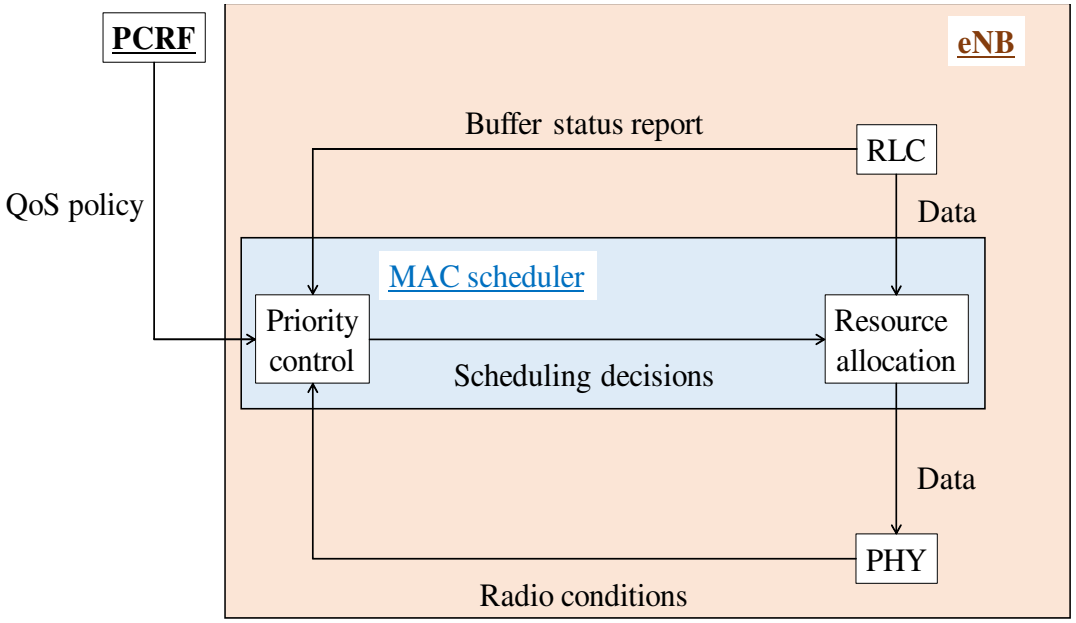

Figure 2.4: eNB components.

fairness, such as Jain's fairness index defined in [50]. A PF scheduler is formulated as

$$P_{pf} = \frac{T_{instant}}{T_{Average}} \tag{2.1}$$

where P_{pf} is the weight of the metric used for user prioritization, $T_{instant}$ is the potentially achievable data rate for a user in the present time slot, and $T_{average}$ is the historical average data rate of the user. With a PF scheduler, packets sometimes exceed the deadline because the same amount of bandwidth is assigned to all UEs regardless of their different deadlines and packet sizes. The maximum carrier-to-interference (Max C/I) scheduler [51, 52] assigns radio resources to a UE based solely on radio quality, without considering packet size or deadline. A Max C/I scheduler increases the system throughput on an eNB but decreases Jain's fairness index. Some UEs that have bad radio quality cannot meet the deadline because Max C/I preferentially allocates radio resources to UEs with good radio quality. Zhang et al. also targeted improvement of the system throughput [53]. The flame level scheduler (FLS) applies control theory to the MAC scheduler [54]. An FLS considers the fluctuation of radio

quality not on the upper level of the scheduler but also on the lower level. Therefore, the FLS suffers from a difference in recognition of radio qualities between upper and lower levels when the radio quality intensely fluctuates. These deadline-unaware schedulers do not meet the deadline because they do not consider it.

Deadline-aware schedulers have been implemented in eNBs to achieve low latency. In the earliest deadline first (EDF) algorithm [55], a data stream with a fast-approaching deadline tends to be prioritized over a data stream with a later deadline. Andreozzi et al. proposed a method considering the deadline and radio quality [56]. Another scheduler called the channel-dependent earliest deadline due (CD-EDD) scheduler [57] considers wireless channel quality but does not consider data size. The modified-EDF-proportional fair (M-EDF-PF) scheduler is an extension of the combined EDF and PF algorithms and is both a channel-aware and QoS-aware scheduler [58]. A packet age-based LTE uplink scheduler was proposed to deliver data before a certain delay budget [59]. However, this scheduler requires another MAC control element. When a cellular network consists of multi-vendor eNBs, all eNBs must support the additional MAC control field.

The payload-size and deadline-aware (PayDA) scheduler is an EDF-based scheduler that performs priority control considering not only the remaining time to the deadline but also the remaining data amount for each data stream [60]. According to [60], PayDA is formulated as

$$\omega_i = \frac{1}{((\tau_i - D_{HOL,i}) \cdot \delta_{left,i})} \tag{2.2}$$

where ω_i is the weight of the metric used for user prioritization, τ_i is the deadline, $D_{HOL,i}$ is the head-of-line delay for the i-th user, and $\delta_{left,i}$ is the remaining data amount of the i-th user. The head-of-line delay $D_{HOL,i}$ indicates the duration of stay

of the first packet in a packet queue on part of the transmitter since its generation.

2.3 Bandwidth Assignment Techniques on eNBs and S/P-GWs

This section describes bandwidth assignment techniques on eNBs and S/P-GWs based on LTE QoS parameters. The LTE QoS parameters include QoS class identifier (QCI), allocation and retention priority (ARP), guaranteed bit rate (GBR), and MBR. QoS rules are applied in accordance with PCC procedure in LTE networks [61].

The QCI, whose range is an integer from 1 to 9, indicates nine different QoS performance characteristics of each IP packet. The different network requirements for each QCI are listed in Table 2.2.

Table 2.2: Network requirements for each QCI.

QCI	Resource type	Priority	Packet delay budget	Packet error loss rate	Example services
1	GBR	2	100 ms	10^{-2}	Conversational voice
2		4	150 ms	10^{-3}	Conversational video (live streaming)
3		3	50 ms	10^{-3}	Real-time gaming
4		5	300 ms	10^{-6}	Non-conversational video (buffered streaming)
5	Non-GBR	1	100 ms	10^{-6}	IMS signaling
6		6	300 ms	10^{-6}	Video (buffered streaming and TCP-based)
7		7	100 ms	10^{-3}	Video, video (live streaming), and interactive gaming
8		8	300 ms	10^{-6}	Video (buffered streaming) and TCP-based
9		9			

The QoS performance characteristics consist of resource type (GBR or non-GBR), priority (1 – 9), packet delay budget (50 ms – 300 ms), and packet loss rate (10^{-2} – 10^{-6}).

The ARP indicates the priority level for the allocation and retention of bearers. LTE networks use ARP to decide whether to accept a request to establish a new bearer or reject the request when resources are depletable due to congestion.

MBR and GBR are used to control bit rates of user traffic. MBR indicates the upper limit of bit rates. When LTE networks have no congestion, user traffic can be delivered at most at the specified MBR. In contrast, because GBR is guaranteed by definition, user traffic is always guaranteed with a specified GBR. Therefore, user traffic is delivered at least at the GBR even when LTE networks have congestion.

Generally, user traffic using different services or applications has different QoS requirements. To satisfy the QoS requirements of all user traffic, LTE networks classify user traffic, such as IP flows and IP packets, as different QCIs and control user traffic using parameters for bandwidth assignment, such as MBR and GBR.

Figure 2.5 shows the location of filters and the relationship between service data flows (SDFs), EPS bearers, and IP flows for the uplink direction. An IP flow generated in a UE is mapped to an EPS bearer by a packet filter called a traffic flow template (TFT) in the UE. When packets in the EPS bearer arrive at an S/P-GW through an eNB, they are mapped to an SDF by a packet filter called an SDF template. The rules of the packet filters, such as the TFT and SDF, are a 5-tuple: source IP address, destination address, source IP port number, destination IP port number, and protocol identifier (ID).

Figure 2.6 shows the locations of filters and the relationship between SDFs, EPS bearers, and IP flows for the downlink direction. An IP flow is mapped to an SDF

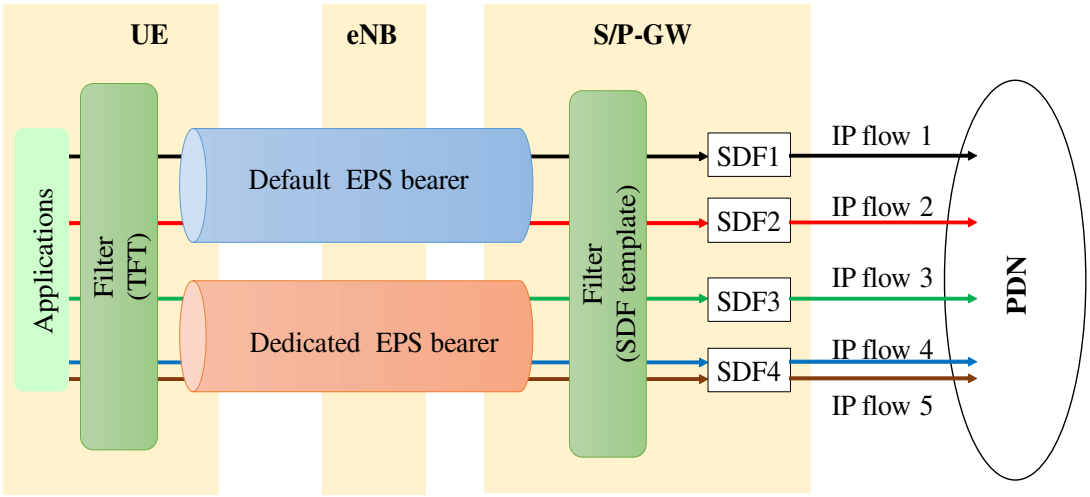

Figure 2.5: Location of filters and relationship between SDFs, EPS bearers, and IP flow for the uplink direction.

whose QoS class of the IP flow matches by an SDF template. SDFs that match a TFT are mapped to an EPS bearer.

MBR is used to control IP flows of two resource types, such as GBR and non-

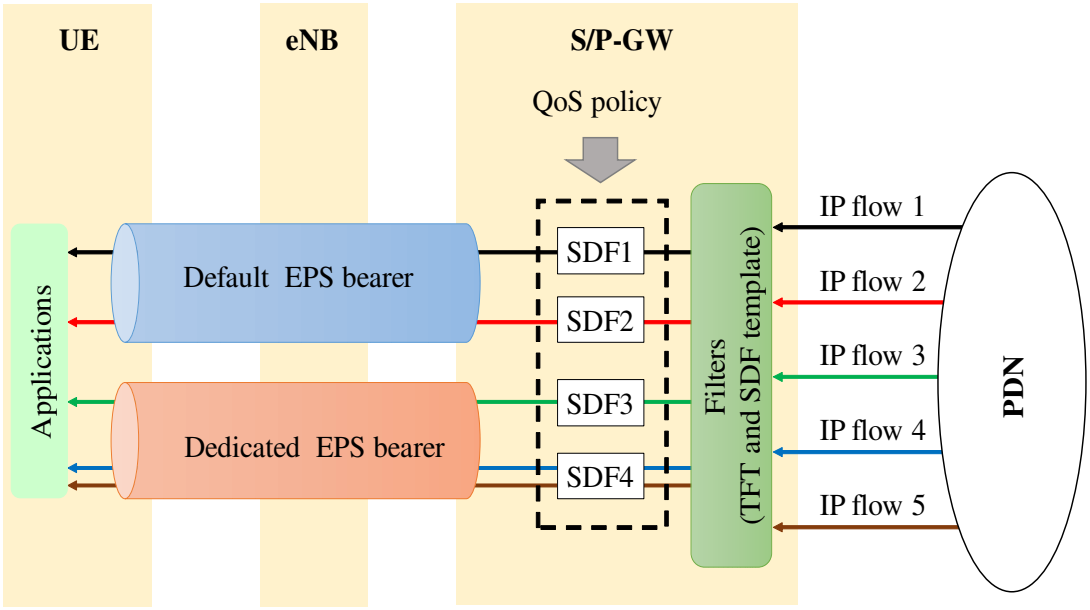

Figure 2.6: Location of filters and relationship between SDFs, EPS bearers, and IP flow for the downlink direction.

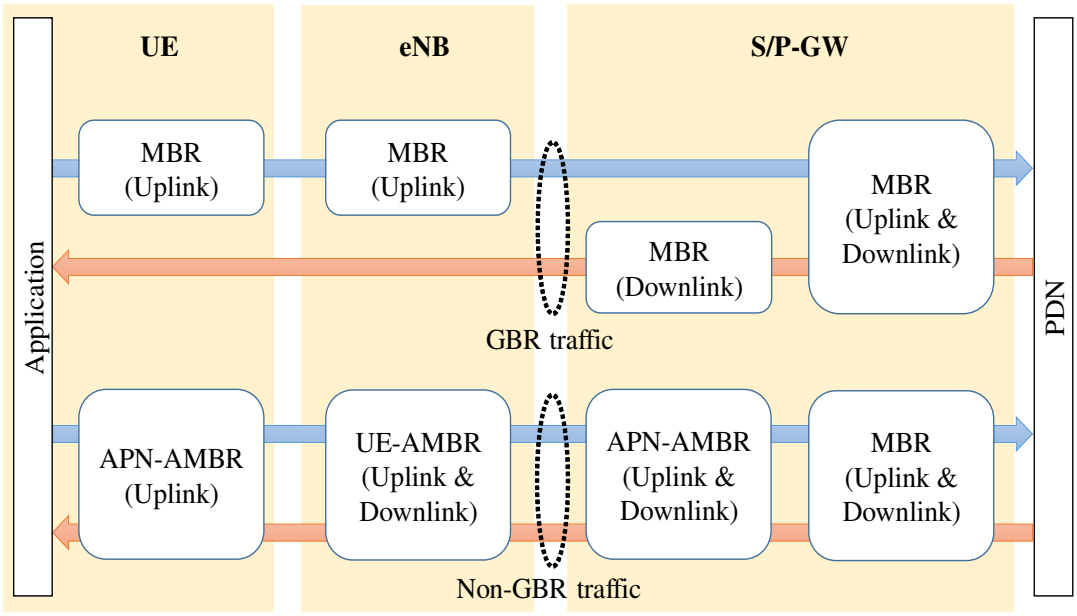

Figure 2.7: Control points of MBR for GBR and non-GBR.

GBR. The control points of MBR for IP flows are shown in Fig. 2.7. For the uplink direction of GBR traffic, MBR controls are performed on UEs, eNBs, and S/P-GWs. When arrival data rate to three nodes, such as UEs, eNBs, and S/P-GWs, exceeds the specified MBR, each node discards packets exceeding the specified MBR and sends the remaining packets in the uplink direction. For the downlink direction of GBR traffic, MBR controls are performed on only S/P-GWs. An S/P-GW has two steps for MBR control. The first MBR control is performed against each SDF. The second MBR control is performed against each EPS bearer.

For non-GBR traffic, there are two types of aggregate MBR (AMBR) control: access point name AMBR (APN-AMBR) and UE-AMBR. The APN-AMBR is the maximum bandwidth that can be shared by all non-GBR bearers in a PDN network. The UE-AMBR is the maximum bandwidth that can be shared in a UE. A UE can be connected to more than one PDN network. In this case, the total APN-AMBR of all PDN networks cannot exceed the UE-AMBR.

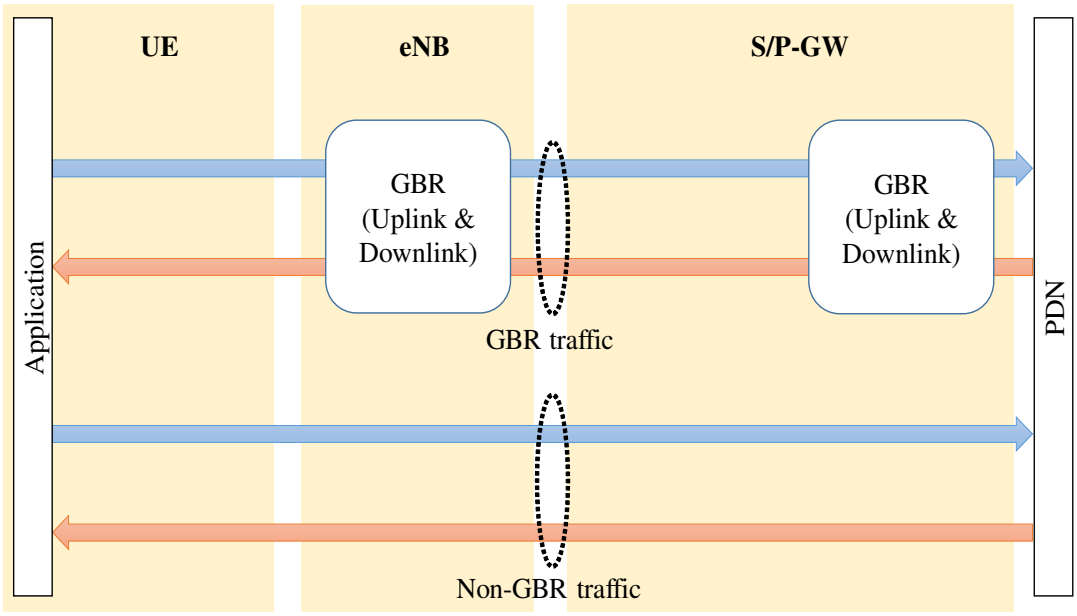

Figure 2.8: Control points of GBR.

MBR has been used to maintain communication quality. Ramamurthi et al. [62] proposed an MBR control method to enhance the quality of experience (QoE) of video streaming, i.e., video QoE. For example, when the video QoE was low, the MBR was set to a larger value to improve the video QoE, allowing a streaming server can provide high-definition video streaming services.

Control points of GBR policing for IP flows are shown in Fig. 2.8. As shown in Fig. 2.8, GBR policing is used to control IP flows whose resource type is GBR only. There are control points of GBR policing on eNBs and S/P-GWs. The granularity of the GBR control is an IP flow. When IP flows for GBR pass through an eNB or S/P-GW in the uplink and downlink directions, the eNB or S/P-GW must guarantee the data rate specified by GBR policing.

Chapter 3

Deadline-Aware Scheduling with Priority Control

3.1 Background and Outline

This chapter focuses on the connected car application described in 1.1.2. LTE networks are assumed to be used as the platform to exchange data for controlling connected cars because LTE networks have been deployed on nation-wide scales. Although 5G networks allow lower latency than LTE networks, it is assumed that 5G networks are partially deployed.

This research assumes LTE V2X via a multi-access edge computing (MEC) server that is located nearby an eNB. The application is assumed to be deployed in a country that drives on the left side, such as Japan. This research investigated an intersection reported in [63] where road traffic collisions often occur. The results show that most road traffic collisions in Japan, where drivers drive on the left side of the road, are caused at intersections: right-turn collisions, rear-end collisions, etc. Right-turn collisions are particularly likely to occur at intersections that are bent sharply to the right, as shown in Fig. 3.1. In countries where drivers drive on the right side of the

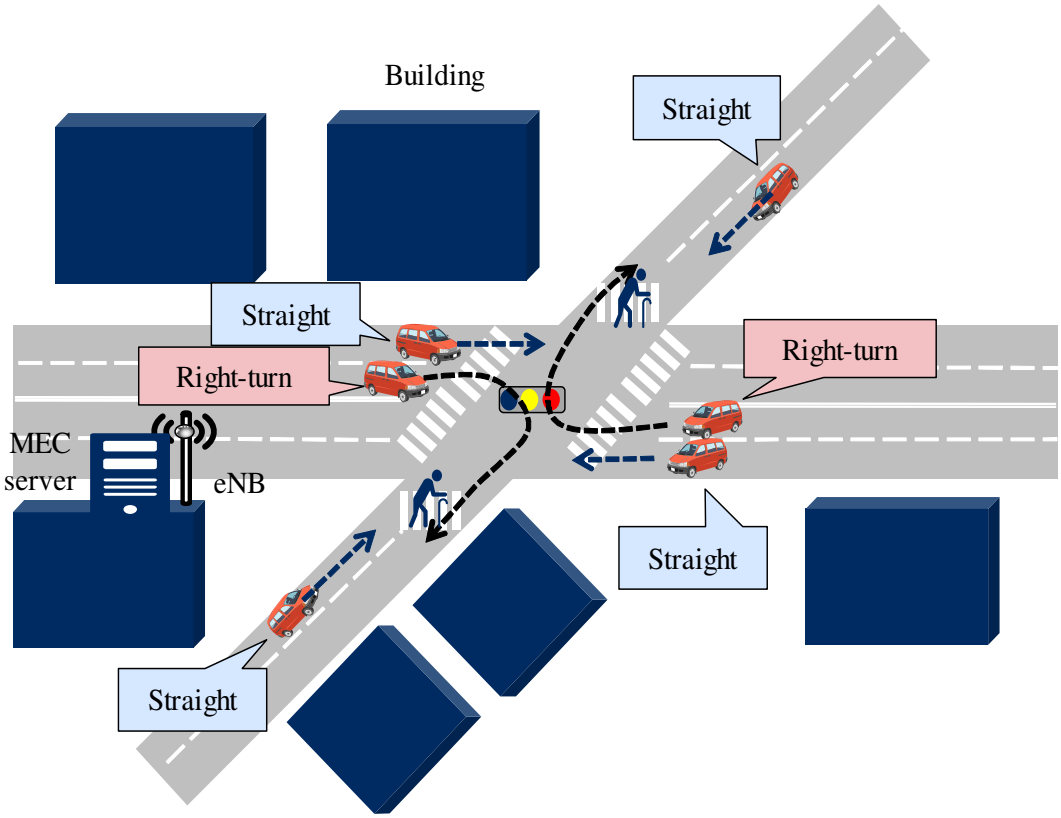

Figure 3.1: Use case intersection.

road, left-turn collisions are more dangerous than right-turn collisions.

This chapter advances discussions on the assumption of a country where drivers drive on the left side of the road. The use case examined in this research is the intersection shown in Fig. 3.1. Vehicles send location information to the MEC server periodically. Then, the MEC server sends a warning message to the same vehicle on the basis of the monitoring image and location information.

The goal of this research is to increase the number of connected users satisfying the QoS requirements. A UE using connected car services sends a request message including location information to an application server and receives a reply message for controlling the services from the application server. In this chapter, a flow that consists of two messages for request and reply is defined as the cycle flow shown in Fig. 1.6, and the deadline of the cycle flow is called the cycle deadline. The cycle

deadline is defined for each data block, which is called a chunk in this dissertation. 3GPP stipulates that the cycle flow should be delivered within a certain time constraint (100 ms) to achieve this use case [40]. Therefore, the cycle deadline is considered to be 100 ms in this research.

The total number of chunks that meet the cycle deadline must be increased to support connected car services on the platform. Goodput (bps) is defined as the total number of chunks that meet deadline constraints per second. System throughput (S_t) is formulated as

$$S_t = \frac{\sum_{k=1}^{N} C_k}{T} \tag{3.1}$$

$$\mathbb{A} = \{k \mid \text{ID of received chunks within } T\}$$

$$|\mathbb{A}| = N$$

where T is observation time, C_k is the size of the k-th chunk, and N is the total number of received chunks. Goodput (S_g) is formulated as

$$S_g = \frac{\sum_{k'=1}^{N'} C_k}{T}, \mathbb{A}' \in \mathbb{A} \tag{3.2}$$

$$\mathbb{A}' = \{k' \mid \text{ID of chunks received before deadline within } T\}$$

$$|\mathbb{A}| = N'$$

where N' is the total number of received chunks that meet the deadline constraint. The degradation of goodput is caused by that of the system throughput because goodput is a subset of system throughput, as shown in (3.1) and (3.2).

When radio resources are allocated without considering radio quality, system throughput is decreased. The PayDA scheduler causes degradation of goodput because

it allocates radio resources without considering radio quality. The PF scheduler also causes degradation of goodput because it allocates radio resources without considering the deadline.

An MEC server is assumed to be used to realize low latency. This server helps effectively provide low-latency services because it is allocated nearby an eNB. Delay-sensitive applications, such as road traffic collision avoidance, are operated from an MEC server [64–68]. By using an MEC server, the distance between vehicles or cameras and the application server can be shorter than when using a mobile network without one.

It is important to ensure low latency in the radio section between UEs, such as vehicles and cameras, and eNBs. The rapid increase in IoT devices has caused growth in data traffic demands [69], and scarcity in last-one-mile bandwidth remains a huge issue across mobile networks [70].

This chapter proposes a deadline-aware scheduling scheme that guarantees the cycle deadline in cases where the chunk size, deadline, and radio quality heavily fluctuate. It identifies emergency UEs and then monitors the progress of data transmission and preferentially gives transmission rights to UEs after considering the radio quality of each UE, thus avoiding exceeding the deadline and degrading system throughput.

It achieves higher goodput than the PayDA and conventional PF schedulers, which are the most implemented scheduling scheme on eNBs and suppresses degradation of system throughput.

3.2 DAS-QF

This section describes the proposed deadline-aware scheduler. The proposed deadline-aware scheduler considers chunk size, deadline, and radio quality and utilizes these to prioritize users such that the deadline constraints are met. The proposed deadline-aware scheduler with consideration for quality fluctuation is referred to as DAS-QF.

DAS-QF has an uplink and downlink deadline coordination function. This function assigns an uplink and downlink deadline to each link in accordance with the cycle deadline. DAS-QF has a resource allocation method that considers radio quality, chunk size, and deadline along with a resource allocation function based on the priority. Its architecture, coordination function, and resource allocation method are presented in 3.2.1, 3.2.2, and 3.2.3, respectively.

3.2.1 Architecture

The architecture of DAS-QF is shown in Fig. 3.2, where it is assumed that various UEs are connected to mobile networks. The UEs are divided into two groups. Group 1 includes UEs that browse websites and streaming services; these are called best effort UEs (BE-UEs). Group 2 includes UEs for road traffic collision avoidance services, such as connected cameras and vehicles; these are called mission-critical UEs (MC-UEs).

The MEC server plays the role of an application server. A mission-critical service, such as road traffic collision avoidance, is operated in the MEC server located in a nearby eNB to achieve low latency. A BE-UE accesses an application server on the Internet as usual to receive a certain desired service. In addition, the MEC server calculates each of the uplink and downlink deadlines and informs the eNB of the

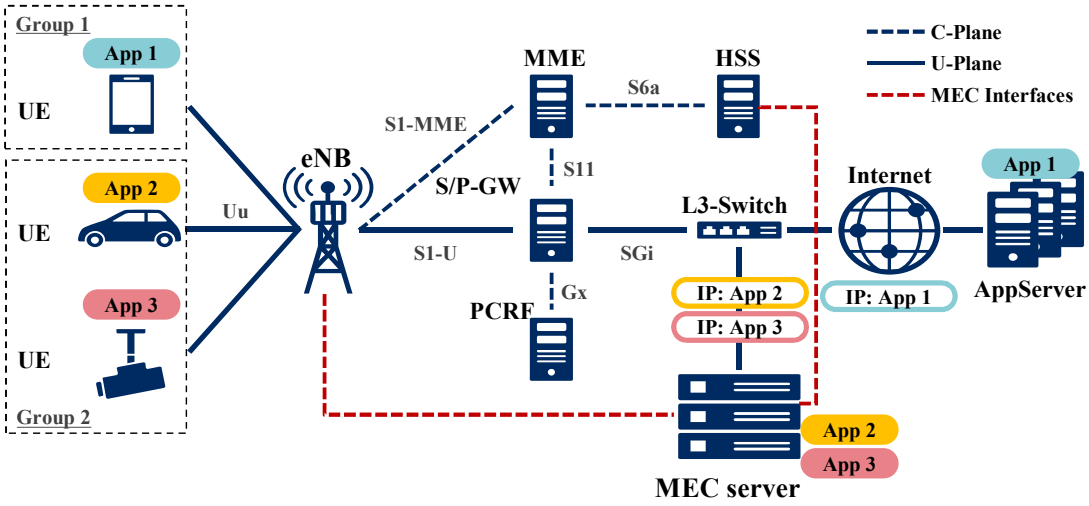

Figure 3.2: Architecture of DAS-QF.

chunk size and calculated deadlines.

The eNB has a MAC scheduling function and decides how to allocate radio resources to UEs. DAS-QF is used as a MAC scheduler in the eNB. It allocates radio resources in accordance with priority, which is calculated based on chunk size, radio quality, and uplink and downlink deadlines.

The HSS has a database for user identification and authentication. It also has a mapping table between a UE and a service that the UE utilizes, and QoS requirements are defined for each service. The HSS has an interface to inform the MEC server of the QoS requirements.

Other devices, such as S/P-GW, MME, and policy, and PCRF, have various functions, as described in Chapter 2.

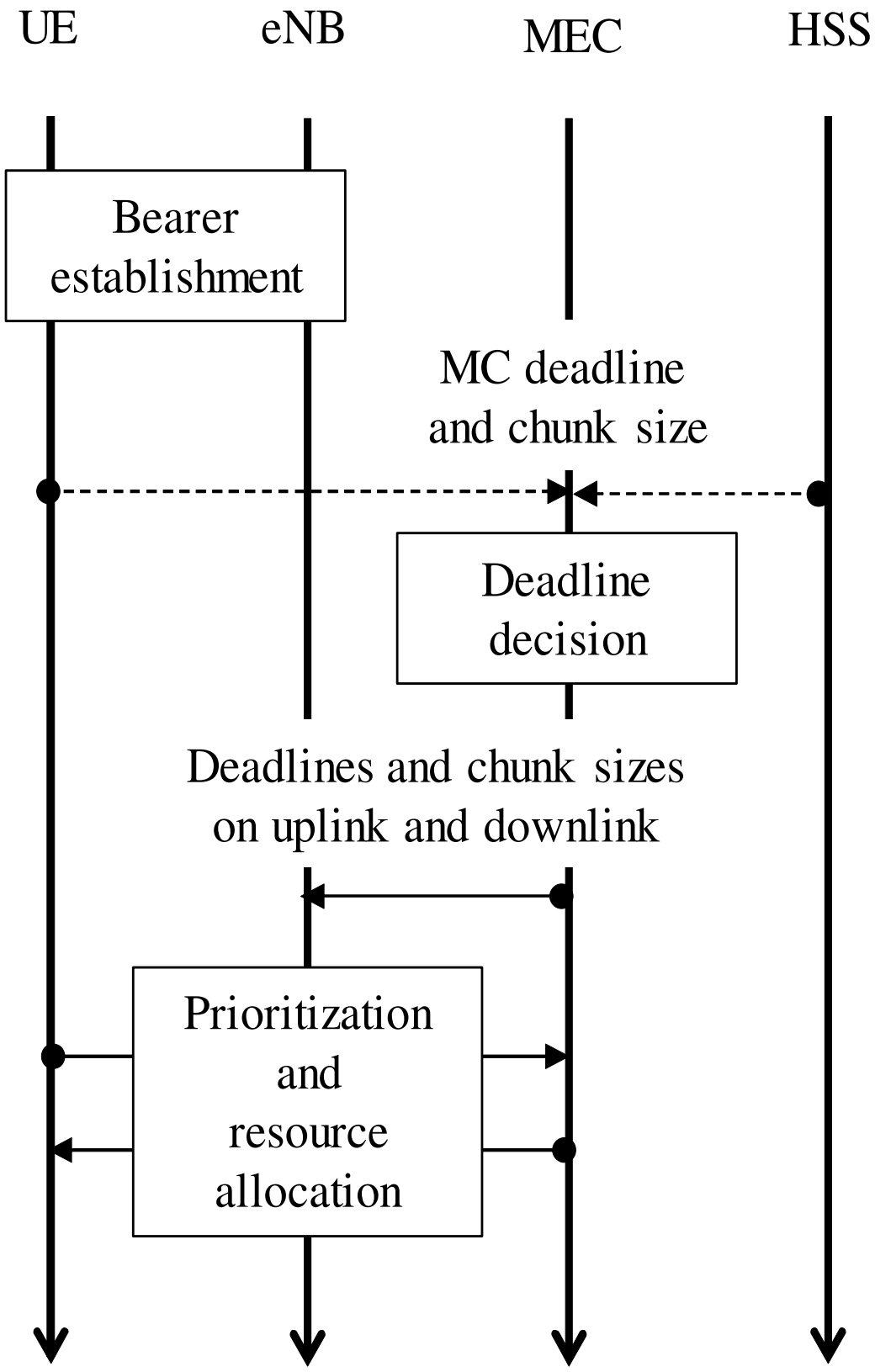

Figure 3.3: Sequence of DAS-QF.

The sequence of DAS-QF is shown in Fig. 3.3. DAS-QF first establishes a bearer between a UE and the eNB. The QoS requirements are defined based on each application and contain the MC deadline. The MEC server receives information on the chunk size via the uplink and the MC deadline via either the source UE or HSS when a new session is established or when the QoS requirements are changed in accordance with application mode switching, i.e., when a vehicle (one of the UEs) sends its real-time location information to avoid collisions in mode 1 and in-vehicle log data, such as a remaining fuel and acceleration work, for performing fuel-saving in mode 2. The deadline of mode 1 is shorter than that of mode 2 since the urgency of mode 1 is higher.

The HSS finds the QoS requirements by referencing the mapping table between a UE and a service that the UE utilizes. The MEC server decides the uplink and downlink deadlines and notifies the eNB of them along with each of the uplink and downlink chunk sizes. It is assumed that the MEC server can understand chunk size on the downlink, as it is created by the MEC server. The eNB calculates UE priority in accordance with each of the uplink and downlink deadlines, each of the uplink and downlink chunk sizes, and radio quality, and it allocates radio resources to the UE based on the calculated priority.

The architecture of DAS-QF shown in Fig. 3.2 can be applied to the use case of connected car services in 5G as well as LTE. In particular, DAS-QF can be applied to both LTE and 5G networks because the objective is to suppress end-to-end delay, including that observed in application layers under a deadline when connected car services are employed.

Algorithm 1 Update of DL deadline

$$ULDeadline \Leftarrow \frac{Deadline - ComputationTime}{2}$$

$$DLDeadline \Leftarrow max(Deadline - ComputationTime - ElapsedTime, Th_1)$$

3.2.2 Update of DL Deadline on MEC Server

This function updates the DL deadline based on the elapsed time, cycle deadline, and computation time of the MEC server. DAS-QF must decide the DL deadline by considering the calculation time of this algorithm. Hence, the calculation time of the algorithm is defined as $ComputationTime$. Algorithm 1 is pseudocode for the DL deadline update. The MEC server sets the UL deadline ($ULDeadline$) to $\frac{Deadline - ComputationTime}{2}$. When the MEC server receives all packets of the chunk from the source UE, it calculates the remaining time to the cycle deadline ($Deadline$) for the chunk—this is done with respect to the cycle deadline ($Deadline$) and elapsed time ($ElapsedTime$)—and sets $DLDeadline$ to the maximum of remaining time ($Deadline - ComputationTime - ElapsedTime$) and Th_1. Th_1 is used to avoid $DLDeadline$ becoming a negative value. Then, the MEC server passes that value as the DL deadline to the eNB connected to the destination UE.

3.2.3 Resource Scheduling Method on eNBs

DAS-QF prioritizes each chunk in accordance with the radio quality between the UE and eNB, the uplink and downlink delay budget, and the chunk size advertised by the MEC server, and then it delivers each chunk to a destination UE within the cycle deadline constraint. The source and destination eNB preferentially allocate radio resources to a high priority UE.

Priority Calculation Algorithm

DAS-QF decides the priority of MC-UE and BE-UE chunks. It distinguishes properties of a chunk (MC-UE or BE-UE) in accordance with bearer information. DAS-QF decides the priority for an MC-UE chunk based on radio quality, chunk size, and deadline to obtain a higher arrival ratio within the cycle deadline. DAS-QF locates emergency UEs by monitoring the progress of data transmission and preferentially gives transmission rights to such UEs to avoid exceeding the deadline. Priority for a BE-UE is decided by P_{pf} in (2.1)

The working concept of DAS-QF is illustrated in Fig. 3.4. DAS-QF judges whether a chunk of an MC-UE is an emergency or non-emergency by comparing the target throughput and requested throughput. The judgement is conducted over one transmission time interval (TTI), which is the same as the scheduling interval. Whether the state of a chunk is emergency is determined not by chunk context but by the progress of data transmission. Considering ease of implementation, a small number of states is desirable because the increase of states causes more complexity and cost increases. Therefore, DAS-QF has only two states: emergency and non-emergency. The target throughput is the average throughput needed to meet the deadline. The requested throughput is the throughput needed to meet the deadline at a certain point in time. If the requested throughput is higher than the target throughput, DAS-QF categorizes the chunk state as an emergency and sets the chunk priority as high to shorten the delay. Otherwise, DAS-QF categorizes the chunk state as non-emergency and sets the chunk priority to low.

The details of DAS-QF are described by Algorithm 2. Algorithm 2 outputs the chunk state, i.e., emergency or non-emergency, and chunk priority. The parameters

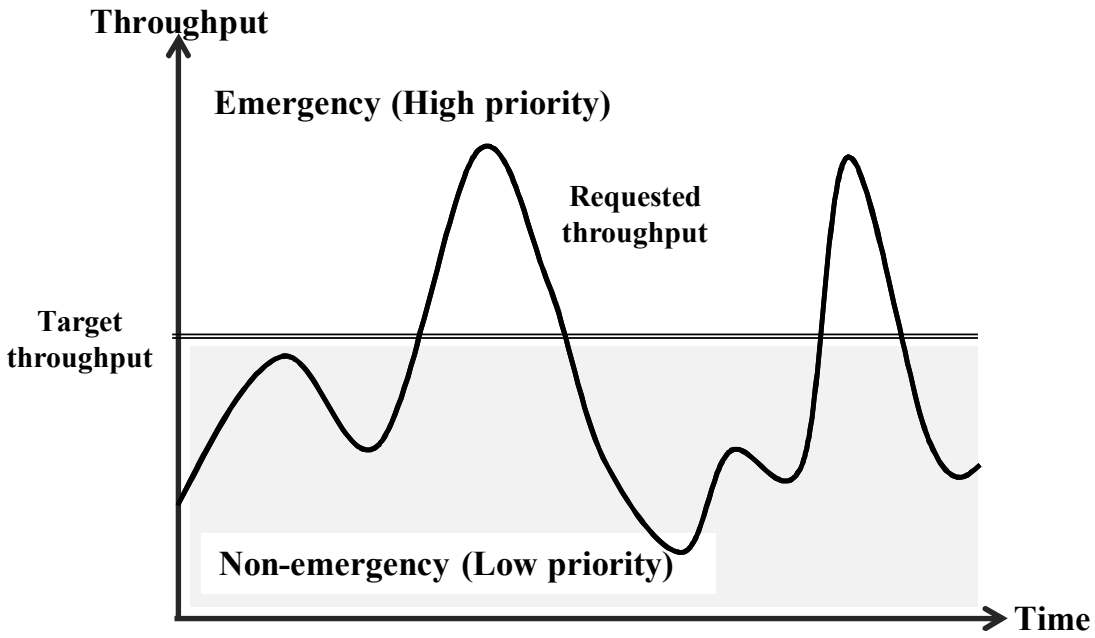

Figure 3.4: Relation between chunk states and priority.

are determined based on the inputs of Algorithm 2, i.e., the deadline, elapsed time, chunk size, instant throughput, average throughput, and radio quality index. Table 3.1 lists the parameters of Algorithm 2 for chunk i at the t-th TTI.

Algorithm 2 Priority calculation algorithm

switch (UE)

case MC-UE**:**

if uplink section **then**

$SubDeadline(i) \Leftarrow ULDeadline(i)$

end if

if downlink section **then**

$SubDeadline(i) \Leftarrow DLDeadline(i)$

end if

$RemainingTime(i,t)$

$\Leftarrow max(SubDeadline(i) - SubElapsedTime(i,t), Th_2)$

$TargetThroughput(i) \Leftarrow \frac{ChunkSize(i)}{SubDeadline(i)}$

$RequestedThroughput(i,t) \Leftarrow \frac{RemainingChunkSize(i,t)}{RemainingTime(i,t)} = \frac{ChunkSize(i) - ReceivedByte(i,t)}{RemainingTime(i,t)}$

$PFmetric(i,t) \Leftarrow \frac{InstantThroughput(i,t)}{AverageThroughput(i,t)}$

if $TargetThroughput(i,t) < RequestedThroughput(i,t)$ **then**

$EmergencyFlag(i,t) \Leftarrow$ **true**

$Priority(i,t) \Leftarrow MCSindex(i,t)$

else

$EmergencyFlag(i,t) \Leftarrow$ **false**

$Priority(i,t) \Leftarrow PFmetric(i,t)$

end if

case BE-UE**:**

$EmergencyFlag(i,t) \Leftarrow$ **false**

$Priority(i,t) \Leftarrow PFmetric(i,t)$

end switch

Table 3.1: Algorithm parameters for chunk i at t-th TTI.

Parameter	Explanation
$TargetThroughput(i)$	Target throughput
$ChunkSize(i)$	Chunk size
$SubDeadline(i)$	Uplink or downlink deadline
$ULDeadline(i)$	Uplink deadline
$DLDeadline(i)$	Downlink deadline
$SubElapsedTime(i)$	Elapsed time on uplink or downlink section
$RequestedThroughput(i,t)$	Requested throughput
$RemainingChunkSize(i,t)$	Remaining chunk size
$ReceivedByte(i,t)$	Received byte
$RemainingTime(i,t)$	Remaining time to deadline
$PFmetric(i,t)$	Priority of chunk i with no emergency flag
$InstantThroughput(i,t)$	Instant throughput of UE with chunk i
$AverageThroughput(i,t)$	Average throughput of UE with chunk i
$EmergencyFlag(i,t)$	Flag to show degree of urgency
$MCSindex(i,t)$	MCS index of UE with chunk i
$Priority(i,t)$	Priority

The priority calculation is executed each TTI. DAS-QF changes the behavior in accordance with UE type (MC-UE or BE-UE). When the UE type is MC-UE, DAS-QF executes the following steps (case MC-UE in Algorithm 2). Th_2 is used to avoid $RemainingTime(i,t)$ being zero or negative. When $RemainingTime(i,t)$ is zero, $RequestedThroughput(i,t)$ is indeterminate.

$RequestedThroughput(i,t)$ has a negative value when $RemainingTime(i,t)$ is negative. If $RequestedThroughput(i,t)$ is higher than $TargetThroughput(i)$, the chunk state is set as "emergency" because it is not likely to meet the deadline. DAS-QF also considers radio quality in addition to the emergency degree to efficiently use radio resources, which are finite. That is, the better the radio quality, the higher the priority. The modulation and coding scheme (MCS) is used as a radio quality index. If $RequestedThroughput(i,t)$ is lower than $TargetThroughput(i)$, the chunk state is set as "non-emergency" because the chunk is making good progress. Then, DAS-QF decides the priority of the chunk using $PFmetric$, which considers the fairness of the transmission right. $PFmetric$ is defined as a fraction of the average throughput and instant throughput.

For a BE-UE (case BE-UE in Algorithm 2), DAS-QF sets the chunk state as "non-emergency" and decides the priority of the chunk using $PFmetric$.

Radio Resource Allocation

After calculating the priority of each chunk with Algorithm 2, DAS-QF allocates radio resources to each chunk in accordance with its priority.

Radio resource allocation assigns radio resources to the UE with the best radio quality among all emergency UEs. By allocating to a good channel UE, the radio resource allocation can improve the goodput while suppressing the degradation of

throughput. If radio resources do not need to be allocated to emergency UEs, the radio resource allocation assigns them to non-emergency UEs in accordance with each $PFmetric$.

3.3 Performance Evaluation

This section compares DAS-QF with the PF and PayDA schedulers. DAS-QF was implemented on ns-3, a general network simulation framework that is widely used in network research [71]. The UEs shown in Fig. 3.2 are defined as vehicles and pedestrians. This research investigated an intersection at which vehicle collisions often occur to evaluate DAS-QF in a realistic environment. The simulation of urban mobility (SUMO) [72, 73] was used to define the mobility model of vehicles and pedestrians. For practical use, DAS-QF is evaluated under the assumption that the network architecture in Fig. 3.2 is LTE. In 5G, it is expected that the performance evaluation results of each method will be improved compared to LTE because the network latency of 5G is smaller. However, DAS-QF is also expected to be effective in 5G because the network may be congested depending on the number of connected cars, and radio quality will fluctuate.

3.3.1 Simulation Environment

A realistic environment was created on ns-3 to investigate an intersection at which vehicle collisions are likely to occur. The results show that most vehicle collisions are caused at intersections: right-turn collisions, rear-end collisions, etc. [33]. Right-turn collisions are particularly likely to occur at intersections that are bent sharply to the right, as shown in Fig. 3.5.

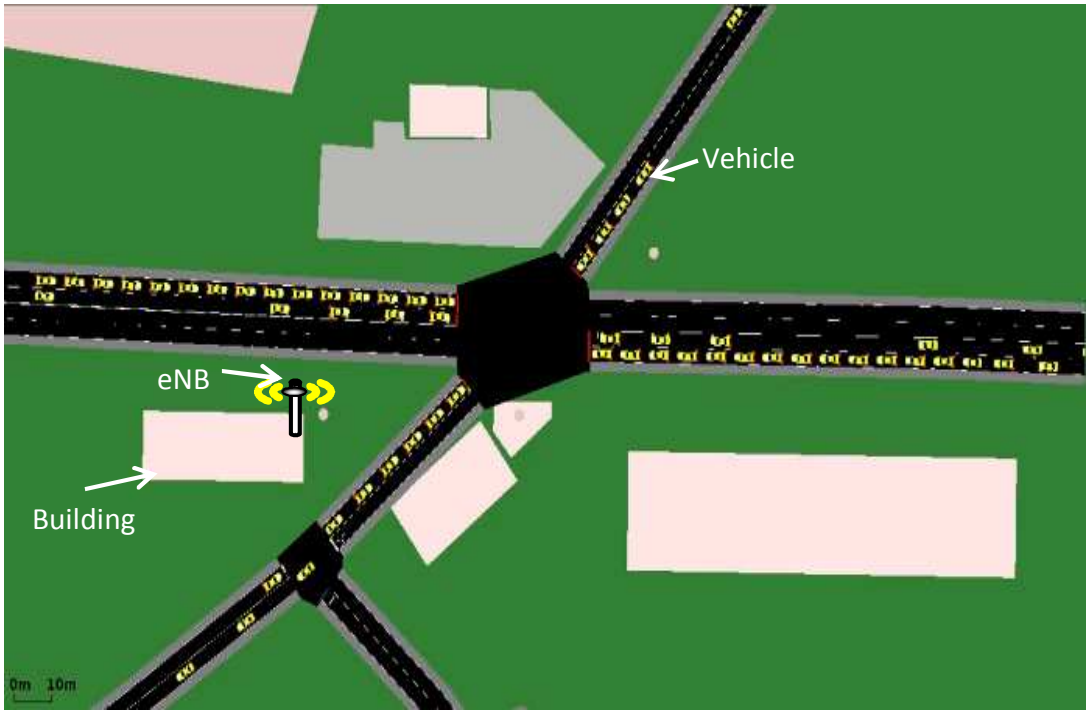

Figure 3.5: Intersection used in the simulation.

The performance evaluation was executed in the environment shown in Fig. 3.5. eNB simulation parameters are listed in Table 3.2. The cell formation was seven hexagonal cells with one eNB allocated at the center of each cell. The road environment was allocated around the intersection on the center cell, which had one eNB and buildings for interference. The six neighboring cells were for interference. Performance on only the center cell was measured. The UL/DL frequency and bandwidth were 2 GHz and 20 MHz, respectively, as used in LTE. Pedestrian A [74] was used as a fading model.

Table 3.3 lists the simulation parameters of the UEs. The MC-UEs (which is delay-sensitive) and BE-UEs (which is delay non-sensitive) were used as vehicles and pedestrians, respectively. The vehicle deadline was 100 ms and was shorter than that of the pedestrian deadline. The numbers of vehicles and pedestrians were 100 each. The first chunk of each vehicle or pedestrian was randomly generated between 0 and 100 ms of simulation time. The next chunks were generated at 100-ms

Table 3.2: Simulation parameters for eNBs.

Number of eNBs	7 (1 cell per eNB)
Radius	289 m (distance between eNBs: 500 m)
UL/DL frequency	2 GHz
Antenna height	32 m
UL/DL bandwidth	20 MHz (=100 resource blocks)
Fading	Pedestrian A

intervals until the end of the simulation. Th_1 and Th_2 were both 1 ms. To evaluate the fundamental performance of DAS-QF, $ComputationTime$, which depends on the machine specifications, was set to 0 ms. Simulation time was 30 s, and the number of simulations was five.

The cycle flow consisted of a UL session and DL session, and the sender and receiver were the same UE. It is assumed that a vehicle sends its own location information, and the MEC server sends an alert message, including location information of neighboring vehicles and pedestrians, to the vehicle. Accordingly, when focusing on a certain vehicle, the sender and receiver are the same. Therefore, DAS-QF was evaluated under the assumption that the sender and receiver are the same. The numbers of cycle flows of MC-UE and BE-UE were both 100. The user data protocol (UDP) was used as a transport protocol to exchange chunks, as 3GPP stipulates that V2X, such as connected car services, uses UDP [75].

Table 3.3: Simulation parameters for UEs.

Tx power	23 dBm
UL power control	Enable
Antenna height	1.5 m
Number of vehicles (MC-UEs)	100
Number of pedestrians (BE-UEs)	100
Traffic generation interval	100 ms
Vehicle deadline	100 ms

3.3.2 Simulation Results of Basic Characteristics

This section presents the simulation results for the environment described in 3.3.1. The goodput of each method was evaluated based on the chunk size of the MC-UE changing as the information amount a vehicle sent or received became greater. Figure 3.6 shows the goodput and application data rate, which is the rate of the amount of application data generated within the simulation time. When the goodput is equal to the application data rate, it means the scheduler is efficient.

As shown in Fig. 3.6, DAS-QF was more efficient than the PF and PayDA schedulers. One reason for this is that DAS-QF preferentially allocates radio resources to chunks that are not likely to meet their deadlines in order to accelerate the transmission progress of such chunks. Therefore, these chunks can be delivered before their deadlines.

When the chunk size was larger than 3 KB, the performances of the PF and

PayDA schedulers degraded. The PF scheduler equally allocates radio resources to each UE without considering deadlines for chunks or chunk size, and such equable radio resource allocation has the potential to increase the number of UEs that fail to deliver a chunk within the deadline. In particular, the performance of the PF scheduler suddenly degraded when the throughput was required to satisfy a cycle deadline constraint that was higher than the throughput equally allocated to each UE. The PayDA scheduler allocates radio resources to whichever UE has little remaining time and a small remaining data amount. The larger the chunk size, the larger the average remaining size. Therefore, the probability of exceeding the deadline increases in proportion to the chunk size.

The ratio of goodput to throughput is formulated as

$$P_A = \frac{S_g}{S_t} \tag{3.3}$$

where S_t and S_g are defined in (3.1) and (3.2), respectively. Figure 3.7 shows P_A from (3.3), i.e., the ratio of goodput (S_g) to throughput (S_t), when using the DAS-QF, PF, and PayDA schedulers.

Degradation of P_A means that finite radio resources were used for chunks that did not meet the deadline even though the scheduler allocated radio resources to MC-UEs, leading to inefficient transmission. When chunk size was equal to or less than 2.4 KB, each method achieved a high performance. P_A of DAS-QF was approximately equal to that of the PF scheduler because DAS-QF decides each chunk's priority by using $PFmetric$ on a non-emergency state. In contrast, when the chunk size was larger than 2.4 KB, the PF and PayDA schedulers exhibited performance degradation, while DAS-QF maintained a consistently high performance. This is because DAS-QF assigns radio resources to monitor the emergency degree and suppress the degradation

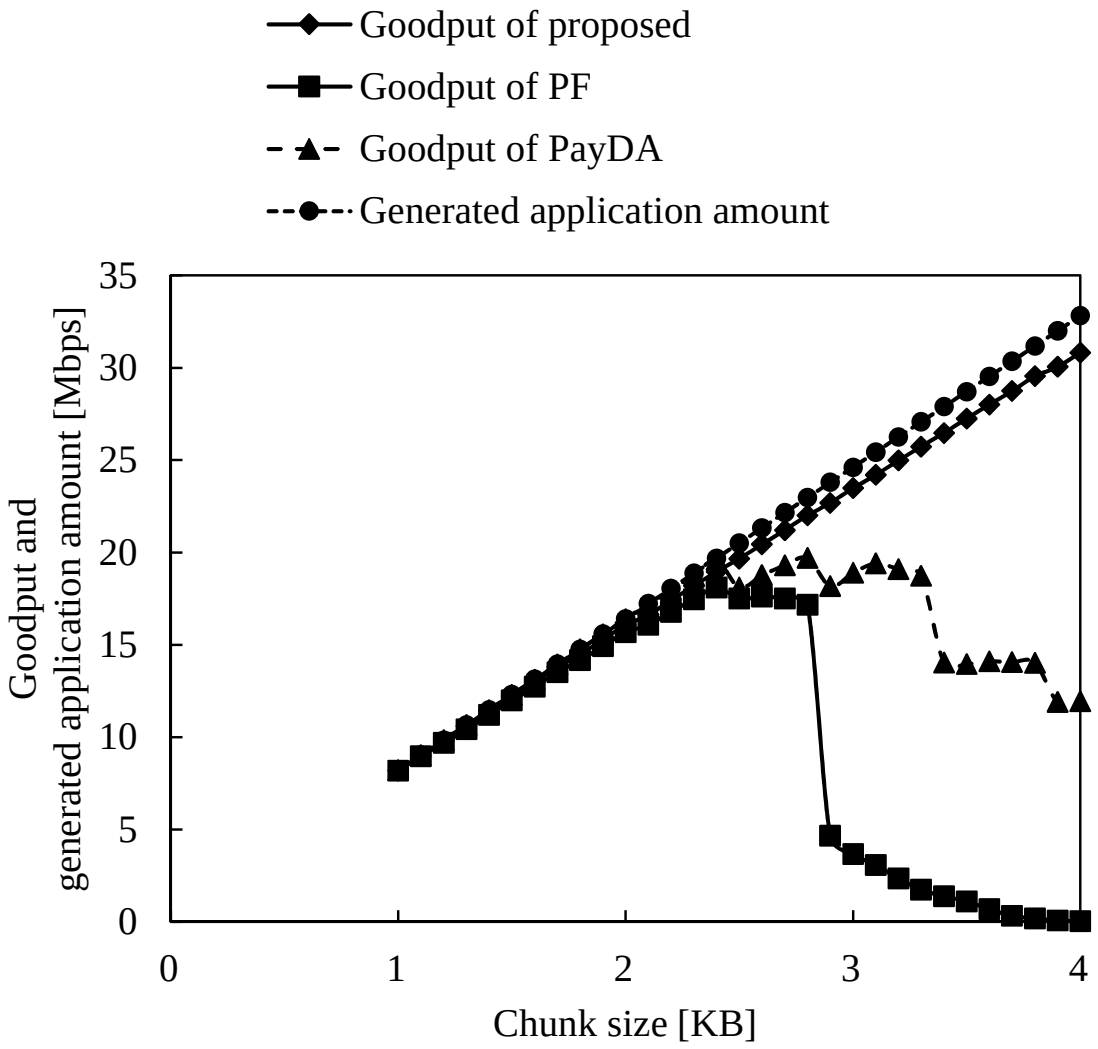

Figure 3.6: Goodput.

of system throughput.

The system throughput of each method is shown in Fig. 3.8; the system throughput is the sum of the total throughput of all MC-UEs and all BE-UEs. The total throughput of all MC-UEs, that of all BE-UEs, and the total goodput are respectively expressed as MC (ST), BE, and MC (GP) in Fig. 3.8.

When the chunk size was 1 or 2 KB, the system throughput of DAS-QF was approximately equal to that of the PF scheduler, as our method prioritizes each chunk with $PFmetric$ in the non-emergency state. In contrast, the system throughput of the PayDA scheduler was lower. This is because the PayDA scheduler has the risk of

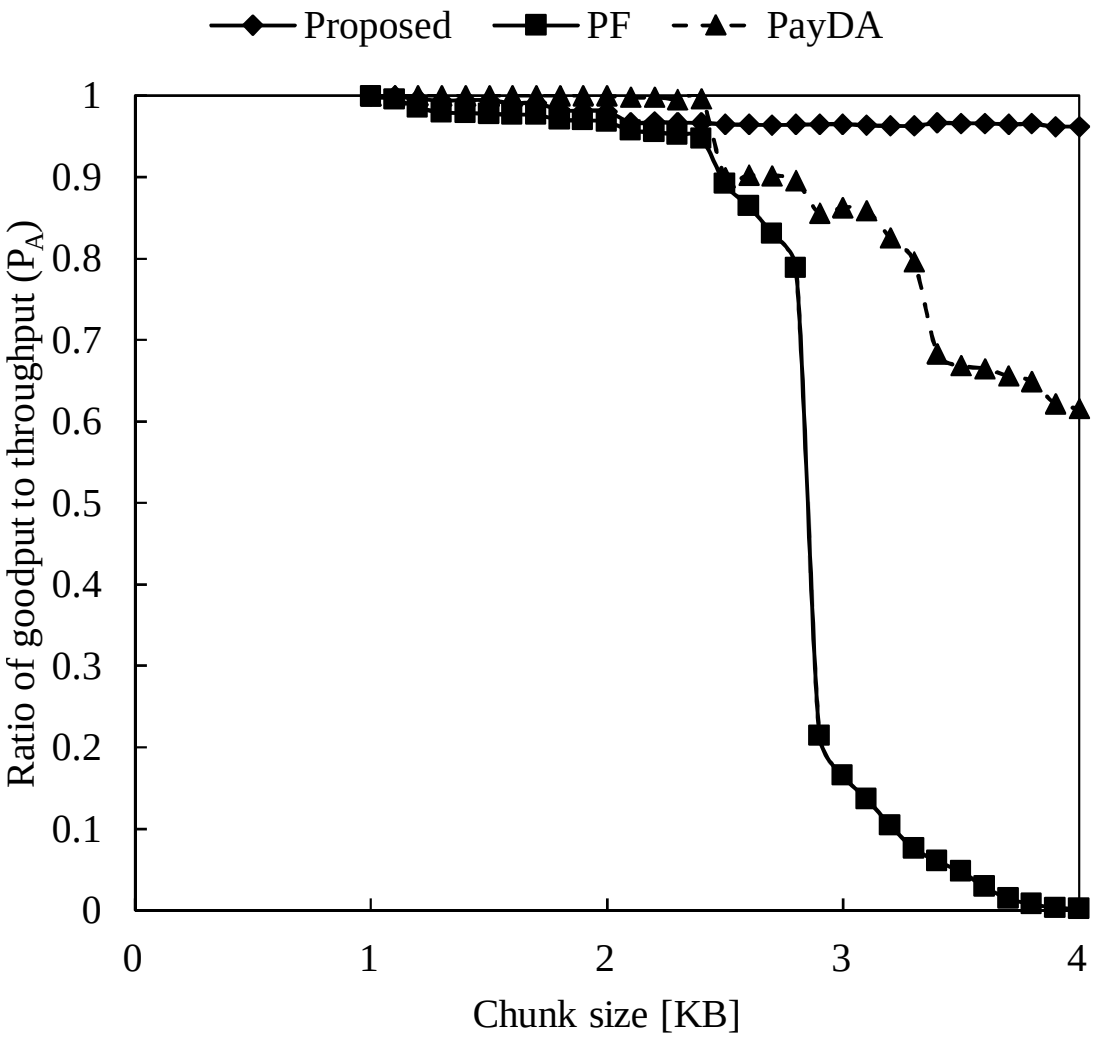

Figure 3.7: Ratio of goodput to throughput.

allocating radio resources with bad radio quality to UEs as it does not consider radio quality.

The PF scheduler had the highest system throughput among all methods when chunk size was 3 or 4 KB. The system throughput of the PayDA scheduler decreased, again because it does not consider the radio quality at all, while DAS-QF obtained a higher system throughput than the PayDA scheduler because it allocates radio resources to UEs that have relatively better radio channel quality. DAS-QF obtained a higher sum of MC (GP) and BE than the PF scheduler. The PF scheduler consumes many radio resources due to chunks that do not meet the deadline, as shown in Fig. 3.8.

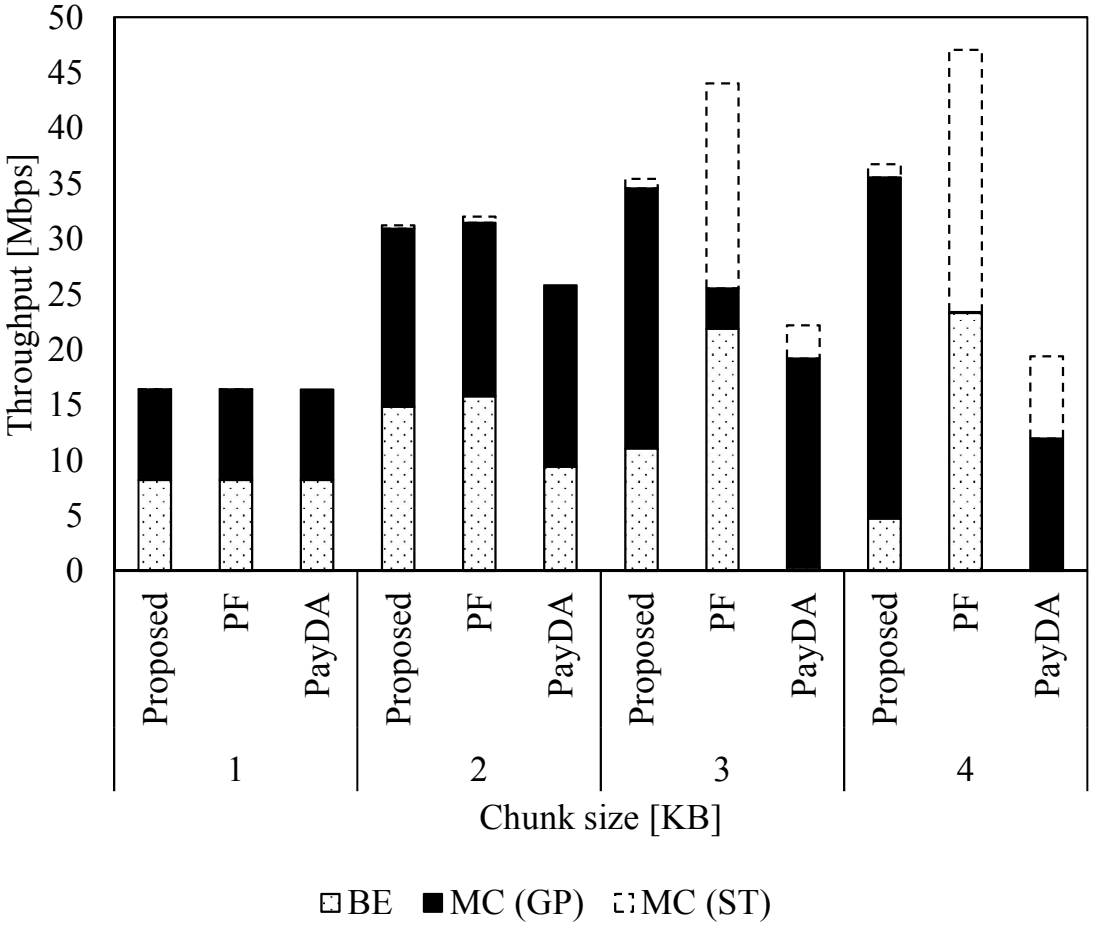

Figure 3.8: System throughput.

With DAS-QF, MC (GP) was approximately equal to MC (ST).

3.3.3 Use Case of Connected Car Services

The parameters for connected car services in terms of delay budget and arrival ratio within the delay budget were presented in [40]. According to [40], in the case of urban intersections, the delay budget and arrival ratio within the delay budget are 100 ms and 95 %, respectively. These are defined as network requirements for safe driving.

This subsection evaluates each method in terms of the number of vehicles satisfying the network requirements for safe driving. One chunk size was set to 3 KB, a relatively large size, because it is assumed that vehicles will be sending camera image data in addition to location and authentication data in the near future.

Figure 3.9 shows the ratio of the number of vehicles satisfying the network requirements to the total number of vehicles when the chunk size was 3 KB. This ratio is formulated as

$$P_b = \frac{N_b}{N_a} \tag{3.4}$$

where N_a is 100 (number of vehicles) and N_b is the number of vehicles satisfying the network requirements. DAS-QF achieved a higher ratio than the PF and PayDA schedulers. In particular, the ratios for DAS-QF and the PayDA scheduler were approximately 0.9 and 0.2, respectively, while the PF scheduler could not satisfy the network requirements for safe driving at all in the simulation environment.

Figure 3.10 shows the success ratio of the chunk of each UE. This ratio is formulated as

$$P_c = \frac{M_b}{M_a} \cdot 100 \tag{3.5}$$

where M_a is the total number of chunks received within simulation time and M_b is the number of chunks received before the deadline within simulation time. The number of MC-UEs is 100, and the horizontal axis indicates UE ID in Fig. 3.10. As can be seen, the success ratios of most UEs were higher than 95% (i.e., the network requirement) with DAS-QF. In contrast, with the PF scheduler, the success ratios of all UEs were distributed between 0% and 40%, and no UEs existed in the area above than 95% because deadline was not considered. The success ratio of the PayDA scheduler was higher than that of the PF scheduler because PayDA considers the remaining time to the deadline and the remaining size. However, its success ratio was lower than that of DAS-QF because it does not consider radio quality. When a scheduler assigns radio resources to UEs with bad radio quality, it leads to inefficient transmission. As a result, the success ratio of the PayDA scheduler decreased.

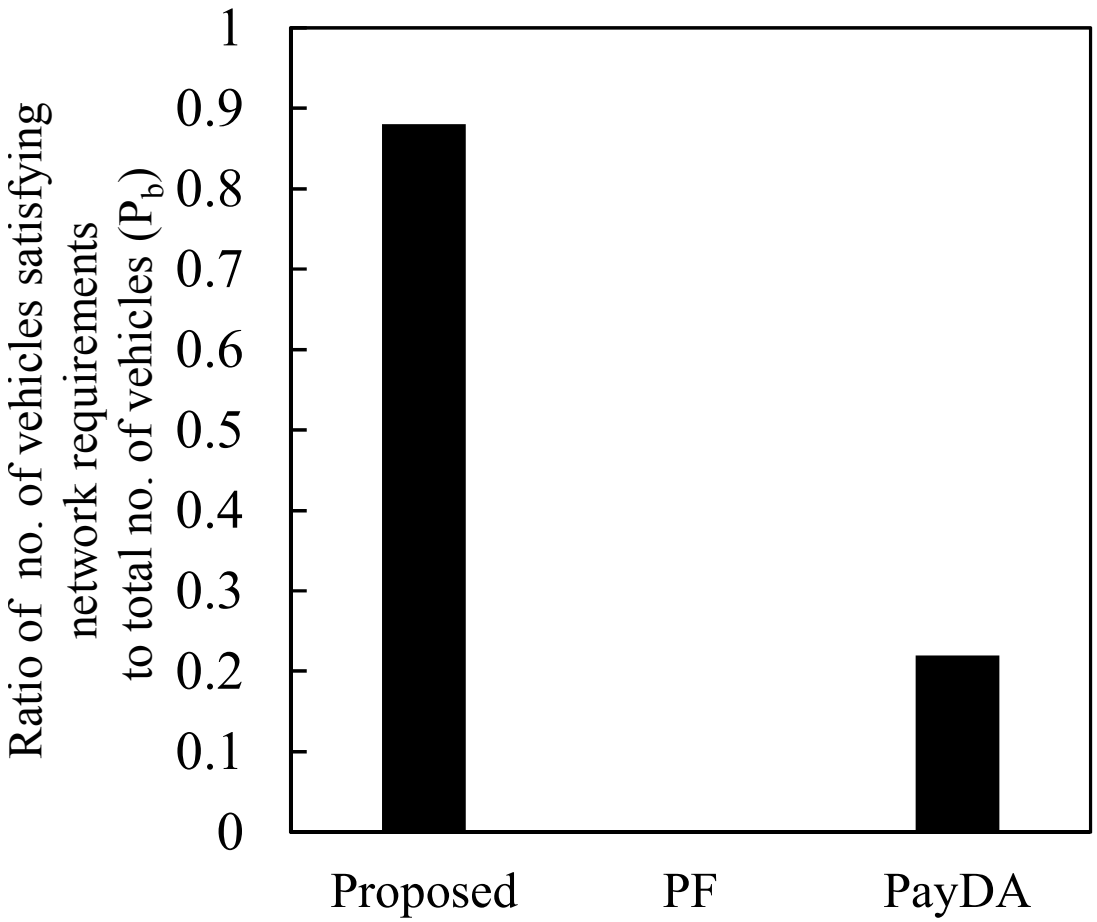

Figure 3.9: Ratio of number of vehicles satisfying network requirements to total number of vehicles.

The reason that DAS-QF can obtain a higher success ratio than the conventional schedulers is that it considers radio quality in addition to chunk size and deadline when allocating radio resources. Table 3.4 lists the average MCS of MC-UEs with DAS-QF in the simulation time. As can be seen, the average MCS of the MC-UEs that had a higher success ratio than 95% and that of the MC-UEs that had a lower success ratio than 95% were 21.1 and 6.2, respectively. This means that DAS-QF can achieve efficient radio resource allocation because it allocates radio resources to UEs with good radio quality. This efficient radio resource allocation enables DAS-QF to achieve a high success ratio compared to those of the PF and PayDA schedulers.

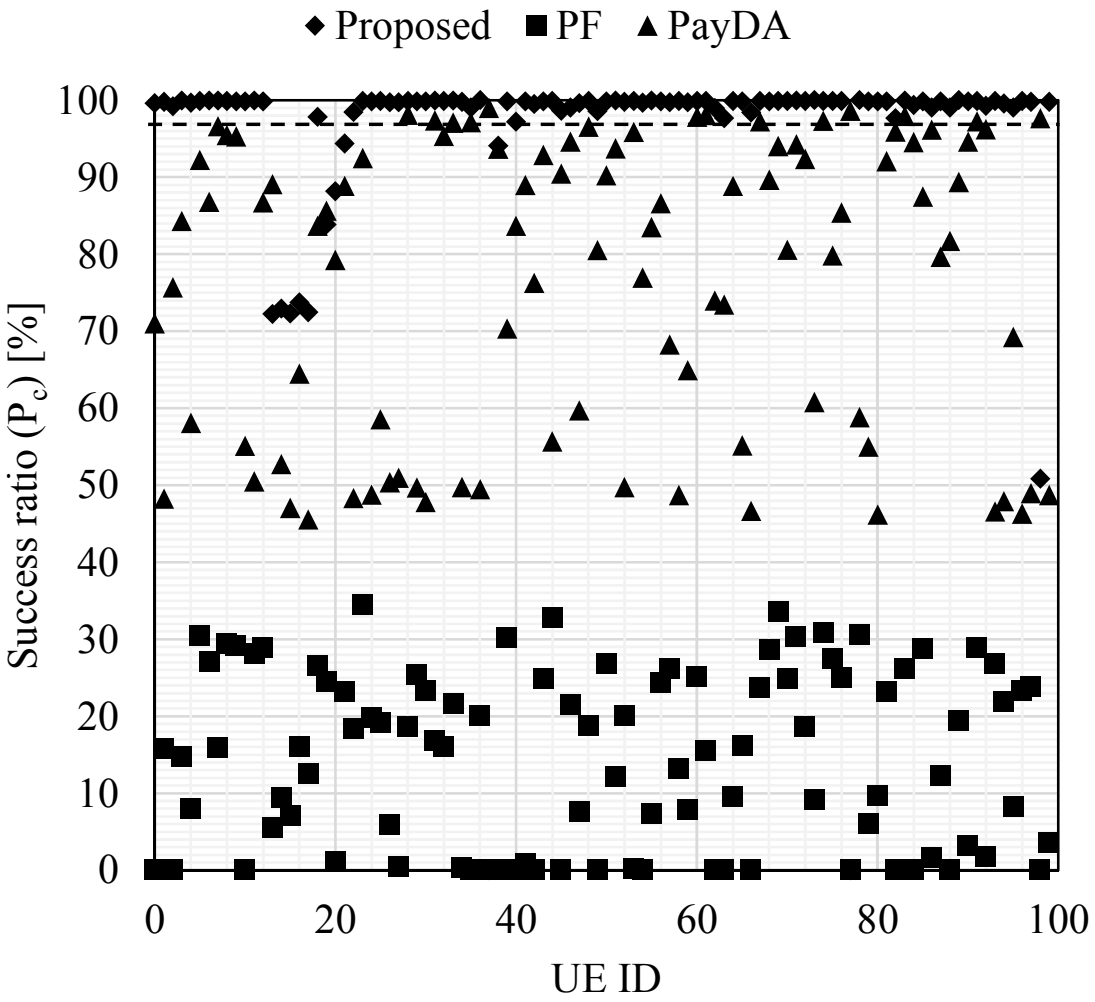

Figure 3.10: Success ratio of each UE.

Table 3.4: Average MCS of MC-UEs with DAS-QF.

UEs (less than 95%)	UEs (more than 95%)
6.2	21.1

3.4 Summary

This chapter presented DAS-QF for connected car services to reduce traffic collisions and evaluated it with ns-3 in a simulation of an intersection at which traffic collisions often occur. This chapter first evaluated the characteristics of chunk size to determine the basic performance when the amount of information that a vehicle sends and

receives increases. The results show that DAS-QF achieved a higher goodput than conventional methods (PF and PayDA schedulers) and was especially efficient when the chunk size was large. This indicates that DAS-QF enables connected cars to send and receive rich content to avoid traffic collisions. DAS-QF was evaluated against network requirements (maximum tolerable delay: 100 ms; arrival ratio within 100 ms: 0.95), and it was found that it could achieve approximately four times the number of vehicles satisfying these requirements than the best-performing conventional method.

Chapter 4

Congestion-Adaptive and Deadline-Aware Scheduling with Priority Control

4.1 Background and Outline

To mitigate road traffic collisions at intersections, it is assumed that connected cameras are allocated around an intersection, as shown in Fig. 4.1. In Fig. 4.1, connected cameras are added to Fig. 3.1.

One of the objectives of this research is to achieve a higher arrival ratio within the deadline for a flow that passes through both the radio uplink and downlink. Ideally, this dissertation envisions a connected camera that periodically sends an image around an intersection to an MEC server. The MEC server then replies with a message for controlling the camera. The message for controlling the camera is assumed to include camera angle, encoding rate, and frame rate. The vehicle periodically sends location information to the MEC server. Then, the MEC server replies to the same vehicle with a warning message for avoiding road traffic collisions. The warning message is generated on the MEC server based on the location information received from the vehicle and the image most recently received from the camera. This dissertation

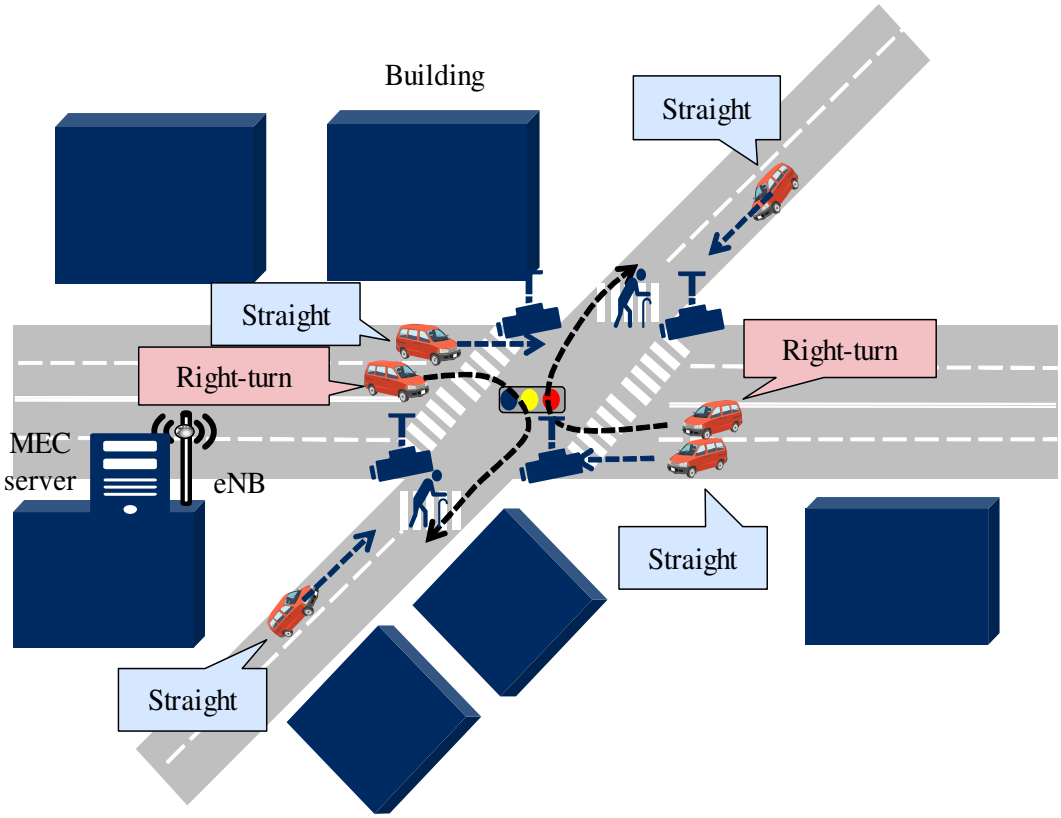

Figure 4.1: Use case intersection with connected cameras.

defines a flow from a vehicle (or camera) to the same vehicle (or camera) via an MEC server, not a flow from a camera to a vehicle via an MEC server, as an MC cycle flow. The whole deadline of the MC cycle flow is called the MC cycle deadline and is defined as 100 ms in this research.

The key to improving the arrival ratio within the MC cycle deadline is determining how to set each of the uplink and downlink deadlines. When deadline exceedance occurs in either the uplink or downlink radio section, the eNB discards the exceeded packet even if does not exceed the MC cycle deadline. This degrades the arrival ratio within the MC cycle deadline. Therefore, it is important to adaptively set each deadline on the uplink and downlink radio sections.

El-Hajj et al. proposed an algorithm to decide the optimal deadlines of the uplink and downlink radio sections [76, 77] based on the throughput in each link. By using

the optimal deadline, the MAC scheduler can guarantee that the first packet of each flow arrives at the destination before the given deadline. However, they did not discuss the effect of heterogeneous traffic environment, i.e., the fact that chunk sizes on the radio uplink and downlink are not always the same on an MC cycle flow. Therefore, this algorithm cannot be applied to the target system of this research.

To improve the arrival ratio within the MC cycle deadline, a deadline coordination function (DCF) for DAS-QF is proposed in this chapter that adaptively allocates uplink and downlink deadlines to each link in accordance with the MC cycle deadline and uplink and downlink congestion levels. Then, DAS-QF with DCF decides priority in accordance with radio quality, chunk size, and allocated uplink and downlink deadlines. Simulation results show that DAS-QF with DCF achieves a higher arrival ratio within the MC cycle deadline than DAS-QF alone.

4.2 Technical Issue with DAS-QF

DAS-QF allocates the same deadline to all UEs. In Fig. 4.2, a vehicle and a camera are connected to the same eNB. The chunk sizes of the vehicle and camera are assumed to be light and heavy, respectively. In this case, the effective uplink delay of the camera chunk is likely to be larger than that of the vehicle chunk. Even so, DAS-QF allocates the same deadline (α), as shown in Fig. 4.2, to both vehicle and camera. When the effective uplink delay of a chunk exceeds the uplink deadline, the chunk is discarded by the eNB. In Fig. 4.2, the effective uplink delays of the vehicle and camera are $\beta_v (< \alpha)$ and $\beta_c (> \alpha)$, respectively. Then, the chunk from the vehicle to the MEC server is not discarded in the eNB, while the chunk from the camera to the MEC server is discarded. As a result, DAS-QF causes degradation of the arrival ratio within the

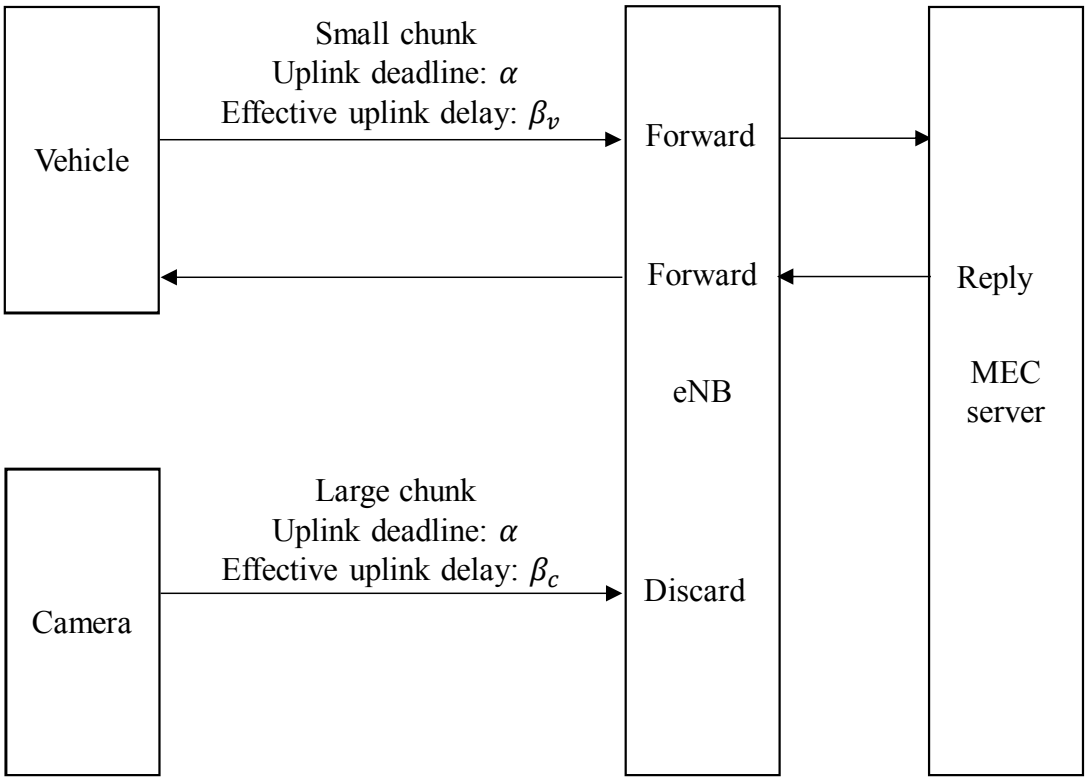

Figure 4.2: Technical issue with DAS-QF.

MC cycle deadline for cameras.

4.3 DCF

This chapter proposes a DCF that decides the uplink and downlink deadlines input to an eNB. The DCF consists of three steps, which are shown in Fig. 4.3. As the first step, the DCF calculates the minimum coordination required to satisfy the arrival ratio within the MC cycle deadline. It calculates the uplink and downlink congestion levels as the second step. As the third step, the DCF updates each of the uplink and downlink deadlines in accordance with the minimum coordination amount and the uplink and downlink congestion levels.

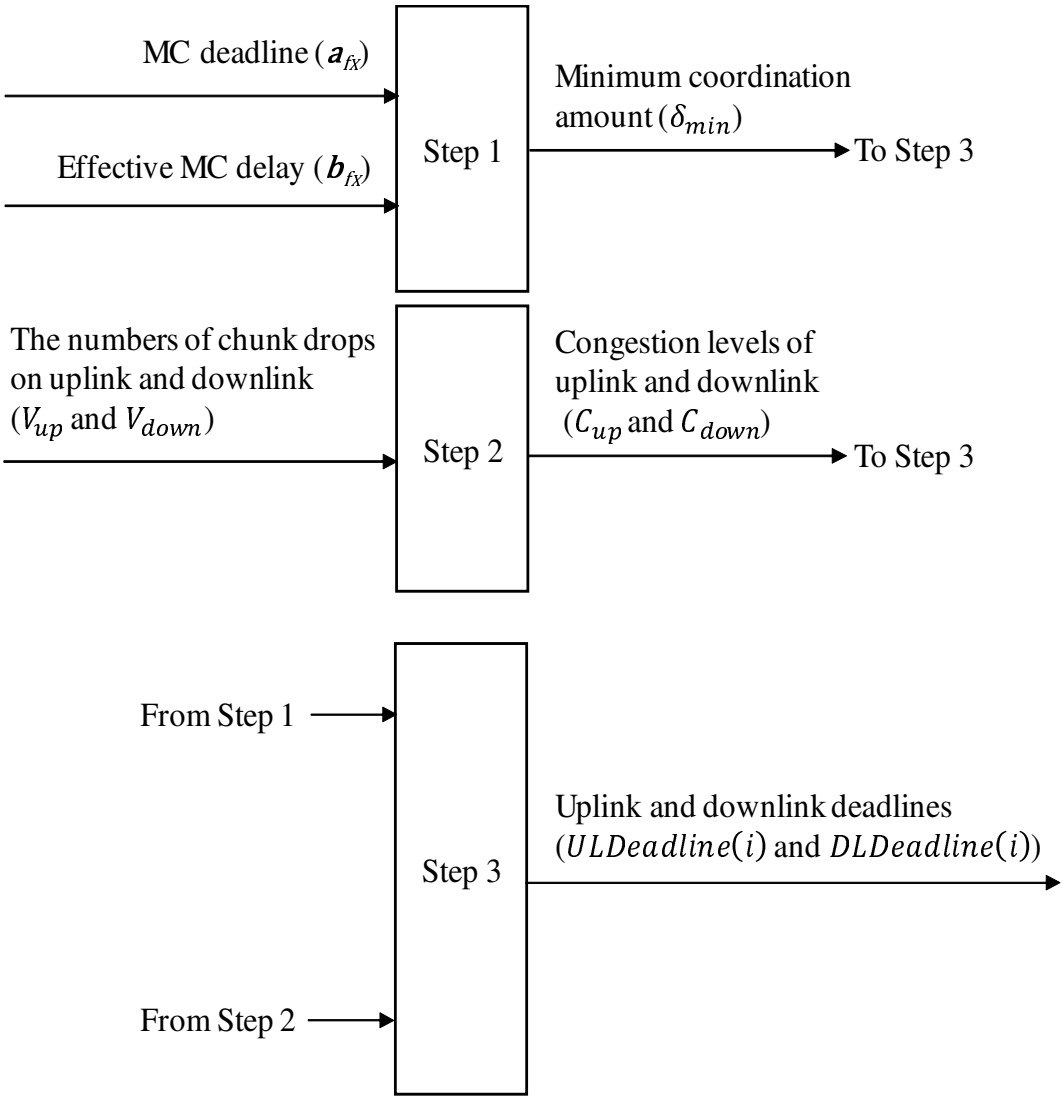

Figure 4.3: Three steps of proposed DCF.

4.3.1 Step 1: Calculation of Minimum Coordination Amount

The number of chunks of MC cycle flow f measured at a certain interval is assumed to be n. This chapter defines n chunks sorted in chronological order as $x_1, x_2, \cdots, x_k, \cdots, x_n$, where $1 \leq k \leq n$. The MC cycle deadline vector $\boldsymbol{a}_{fx}$ is expressed as

$$\boldsymbol{a}_{fx} = (a_{x_1}, a_{x_2}, \cdots, a_{x_k}, \cdots, a_{x_n}), \tag{4.1}$$

where a_{x_k} is the MC cycle deadline of chunk x_k. The effective MC cycle delay vector $\boldsymbol{b}_{fx}$ is expressed as

$$\boldsymbol{b}_{fx} = (b_{x_1}, b_{x_2}, \cdots, b_{x_k}, \cdots, b_{x_n}), \tag{4.2}$$

where b_{x_k} is the effective MC cycle delay of chunk x_k. A UE is assumed to embed a unique ID into a message and records the time when it sends the message to the MEC server. When the MEC server sends a reply message to the same UE, the MEC server also embeds the received unique ID into the reply message. The UE can measure effective MC cycle delay based on the time the message was sent to the MEC server, the time the message was received from the MEC server, and its unique ID. The UE advertises an effective MC cycle delay to the MEC server.

The remaining time vector $\boldsymbol{\delta}_{fx}$ is expressed as

$$\boldsymbol{\delta}_{fx} = (\delta_{x_1}, \delta_{x_2}, \cdots, \delta_{x_k}, \cdots, \delta_{x_n}), \tag{4.3}$$

where δ_{x_k} is the remaining time to the MC cycle deadline of chunk x_k. The remaining time vector $\boldsymbol{\delta}_{fx}$ is formulated as

$$\boldsymbol{\delta}_{fx} = \boldsymbol{a}_{fx} - \boldsymbol{b}_{fx}. \tag{4.4}$$

The DCF sorts $x_1, x_2, \cdots, x_k, \cdots, x_n$ in descending order of remaining time to the MC cycle deadline, where the sorted chunks are represented by $y_1, y_2, \cdots, y_k, \cdots, y_n$. The sorted remaining time vector $\boldsymbol{\delta}_{fy}$ is expressed as

$$\boldsymbol{\delta}_{fy} = (\delta_{y_1}, \delta_{y_2}, \cdots, \delta_{y_k}, \cdots, \delta_{y_n}), \tag{4.5}$$

where δ_{y_k} is the remaining time to the MC cycle deadline of chunk y_k. Then, the minimal coordination amount δ_{min} is expressed as

$$\delta_{min} = \delta_{y_{\lceil n \cdot Q_f \rceil}}, \tag{4.6}$$

where Q_f is the required arrival ratio within the MC cycle deadline of flow f, and $0 < Q_f \leq 1$. When δ_{min} is greater than or equal to zero, the MC cycle flow f satisfies the required arrival ratio within the MC cycle deadline. Therefore, Steps 2 and 3 are not executed when $\delta_{min} \geq 0$. In contrast, when δ_{min} is negative, the MC cycle flow f does not satisfy the required arrival ratio within the MC cycle deadline. To satisfy the required arrival ratio within the MC cycle deadline, Steps 2 and 3 are executed when $\delta_{min} < 0$.

4.3.2 Step 2: Estimation of Congestion Level

In this step, the DCF estimates the uplink and downlink congestion levels in accordance with the number of chunk drops in the uplink and downlink. The uplink congestion level C_{up} is formulated as

$$C_{up} = \frac{V_{up}}{V_{up} + V_{down}}, \tag{4.7}$$

where V_{up} and V_{down} are the numbers of chunk drops at uplink and downlink, respectively. The downlink congestion level C_{down} is formulated as

$$C_{down} = \frac{V_{down}}{V_{up} + V_{down}}. \tag{4.8}$$

V_{up} and V_{down} are measured by the eNB, and the eNB advertises V_{up} and V_{down} to the MEC server. A time stamp is assumed to be used for the measurement of V_{up} and V_{down}, and all clocks of UEs, eNBs, and the MEC server are synchronized. A UE embeds a time stamp into a chunk when it sends the chunk to the MEC server, and the MEC server embeds another time stamp into the chunk when it sends the chunk to the UE. The eNB calculates the remaining time to the deadline using the time stamp and deadline when it received the chunk. The eNB discards the chunk and increases the drop count when the chunk does not meet the deadline.

When the number of chunk drops is high, it is assumed that the congestion level is high compared to when the number of chunk drops is low.

4.3.3 Step 3: Update of Uplink and Downlink Deadlines

This step updates the uplink and downlink deadlines of chunk i in accordance with the minimum coordination amount δ_{min} and uplink and downlink congestion levels C_{up} and C_{down}, respectively. The uplink and downlink deadlines of chunk i are $ULDeadline(i)$ and $DLDeadline(i)$ and are defined by (4.9) and (4.10), respectively.

$$ULDeadline(i) = \delta_{min} \cdot C_{down} + ULDeadline(i-1) \tag{4.9}$$

$$DLDeadline(i) = \delta_{min} \cdot C_{up} + DLDeadline(i-1) \tag{4.10}$$

The uplink and downlink deadlines of the chunk whose effective MC cycle delay exceeds the MC cycle deadline are set to smaller values in the range that is not generating chunk drops in the eNB in order to raise the priority of the chunk. A negative δ_{min} means that the previous chunk $i-1$ could not be delivered within the MC cycle deadline. The next chunk i also cannot meet the MC cycle deadline under the same network conditions as the previous chunk. Therefore, the sum of the next uplink and downlink deadlines is set to a shorter value than the sum of the previous uplink and downlink deadlines. This results in a shorter effective MC cycle delay for the next chunk than that for the previous chunk because the next chunk has priority use of radio resources according to Algorithm 1. By this step, it is expected that the next chunk meets the MC cycle deadline.

After Step 3, the MEC server sends $ULDeadline(i)$ and $DLDeadline(i)$ to the eNB that the UE is connected to. The eNB then decides the UE priority based on

the radio quality, chunk size, and target deadline information received from the MEC server and allocates radio resources to the UE based on this priority.

4.4 Performance Evaluation

This chapter investigates the effect of the DCF through simulations using ns-3. In particular, this chapter compares the performance of DAS-QF with DCF to that with DAS-QF alone. In Chapter 3, it was shown that DAS-QF is more effective than the conventional PayDA and PF schedulers.

4.4.1 Simulation Environment

This research investigated an intersection where road traffic collisions often occur to evaluate the DCF in a realistic environment. In particular, the realistic environment shown in Fig. 4.1 was implemented in ns-3. SUMO [72, 73] was used to define the mobility model of vehicles to simulate the realistic mobility behavior of vehicles as well as the performance evaluation of Chapter 3.

This chapter evaluates DAS-QF with DCF under the LTE assumption, but it is not limited to LTE. For example, it is expected that the 5G mobile network will improve the arrival ratio within the MC cycle deadline compared to with LTE because the network latency of 5G is smaller. DAS-QF with DCF is expected to be effective in 5G because the network may be congested by increasing the number of connected devices.

The simulation parameters of the eNB are the same as those used in the performance evaluation of Chapter 3 and are listed in Table 3.2. The cell formation is seven hexagonal cells with one eNB allocated at the center of each cell. The road environment

is allocated around the intersection on the center cell, which has one eNB and several buildings for interference. The six neighboring cells are also for interference. This chapter measures the performance only on the center cell. A new vehicle is assumed to appear in the center cell at the same time as a vehicle moves out of the center cell range so that the number of vehicles and network load in the center cell remain constant. Therefore, this simulation does not have a handover situation. When a new vehicle appears in the center cell, it is assumed that the new vehicle sends a chunk after bearer establishment by the initial attaching procedure. The effective MC cycle delay does not include the delay for bearer establishment. The uplink and downlink frequency and bandwidth are 2 GHz and 20 MHz, respectively, as used in LTE. Pedestrian A [74] is used as the fading model.

Table 4.1 lists the simulation parameters of the UEs in the use case with connected cameras. MC-UEs, which are delay-sensitive, are vehicles and cameras, and BE-UEs, which are non-delay-sensitive, are pedestrians. A BE-UE accesses an application server on the Internet as usual to receive a certain desired service. A BE-UE sends a chunk to the application server. The application server replies with a chunk for the desired service to the BE-UE. This dissertation calls a flow from a BE-UE to the same BE-UE via an application server the BE cycle flow. The effective delay of the BE cycle flow is defined as the effective BE delay. The number of cameras is 4, assuming that each camera is located around each corner of the intersection. The first chunk of each vehicle, camera, or pedestrian was randomly generated between 0 and 100 ms of simulation time [78]. The next chunks were generated at 100-ms intervals until the end of the simulation.

This chapter investigates a proper bitrate of video traffic with the Toomer, which is a video of a certain intersection and is provided in [79]. This research set the frame

Table 4.1: Simulation parameters for UEs in the use case with connected cameras.

Tx power	23 dBm
Uplink power control	Enabled
Antenna height	1.5 m
Number of vehicles	100
Number of cameras	4
Number of pedestrians	100
Chunk size of vehicle in uplink and downlink	1.7 KB
Chunk size of camera in uplink	37.5 KB
Chunk size of camera in downlink	1.7 KB
Chunk size of pedestrian in uplink and downlink	1.7 KB
Traffic generation interval	100 ms
MC cycle deadline	100 ms

rate, resolution, and codec to 10 fps, 1280*720 (HD 720), and H.264, respectively. Cars and pedestrians were detected with YOLOv3 [80], and the mean average precision (mAP) was measured for cars and pedestrians. When the video was transcoded to 3 Mbps, the mAPs of 95% of all frames were more than 0.9. Assuming that this accuracy was sufficient for practical application, the video traffic was set to 3 Mbps. Because each camera sent an image to the MEC server every 100 ms, the size of each image data chunk was 37.5 KB. In the 3GPP specification, codec delay is not included

in the deadline. Therefore, the codec delay was not included in the effective MC cycle delay.

The sizes of other chunks (such as location information from a vehicle, messages for warning and control from the MEC server, and messages that pedestrians send and receive) were all 1.7 KB. The size of chunk from a vehicle to the MEC server and the size of a chunk from the MEC server to a vehicle were also the same. Therefore, in DAS-QF, the uplink and downlink deadlines were both set to 50 ms. Th_2 was set to 1 ms.

UDP was used as a transport protocol to exchange chunks, as 3GPP stipulates that V2X, such as connected car services, use UDP [75]. The maximum transmission (Tx) power was 23 dBm, and the power control specified in [81] was applied to the simulation environment.

In this research, it was assumed that the delay for the sequence of updating uplink and downlink deadlines was negligible because the MEC server was located near the eNB and notified the eNB of the updated deadlines by using a control plane, whose congestion level is lower than that of a data plane.

4.4.2 Simulation Results

The QoS requirements for connected car services in terms of MC cycle deadline and arrival ratio within the MC cycle deadline were presented in [40]. According to [40], in the case of urban intersections, the MC cycle deadline and arrival ratio within the MC cycle deadline should be 100 ms and 95%, respectively. These were defined as network requirements for connected car services in this research.

Figure 4.4 shows the arrival ratio within the MC cycle deadline, which is formulated

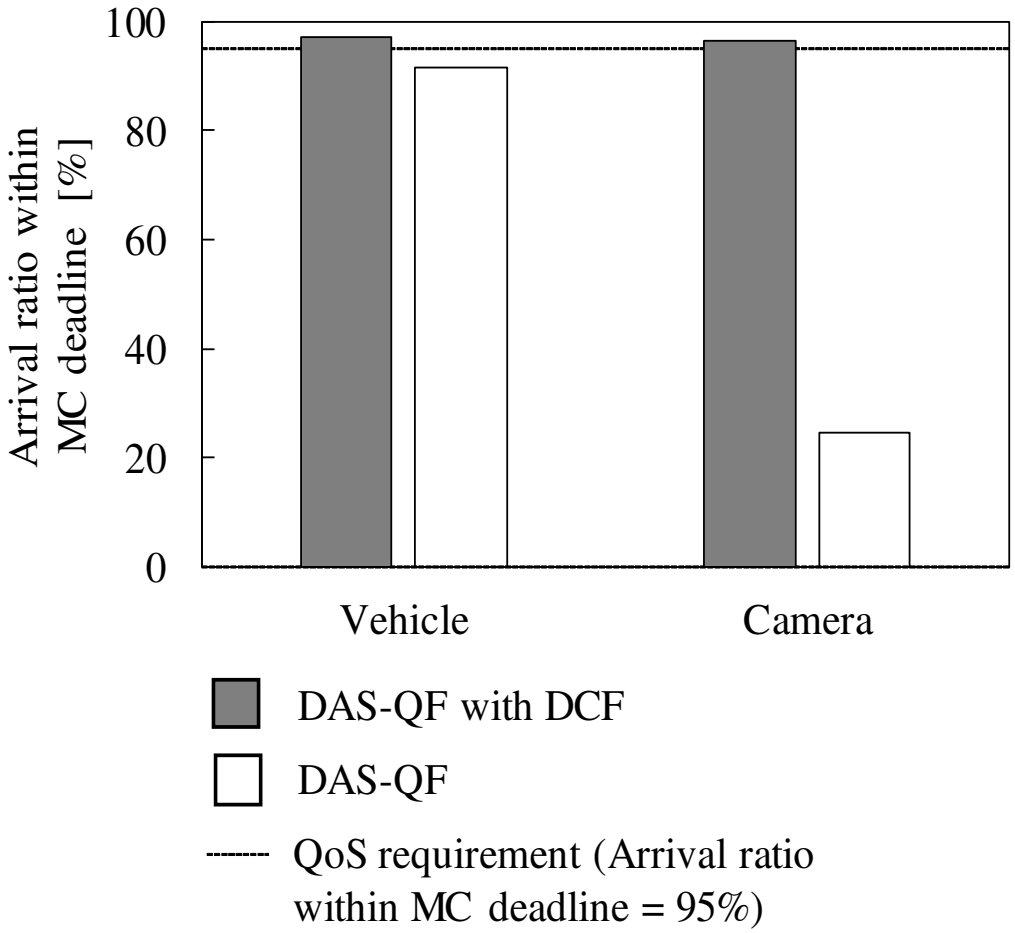

Figure 4.4: Arrival ratio within MC cycle deadline.

as

$$P_d = \frac{X_b}{X_a}, \tag{4.11}$$

where X_a is the total number of chunks that MC-UEs sent within the simulation time and X_b is the number of chunks that MC-UEs received before the MC cycle deadline within the simulation time.

As shown in Fig. 4.4, DAS-QF with DCF obtained higher vehicle and camera arrival ratios within the MC cycle deadlines than DAS-QF alone. Specifically, DAS-QF with DCF could attain 95% of arrival ratio within the MC cycle deadline at both vehicle and camera, while DAS-QF alone could not.

DAS-QF also obtained a high arrival ratio within the MC cycle deadline compared to the camera arrival ratio within the MC cycle deadline. In this chapter, the size of chunk from a vehicle to the MEC server and the size of chunk from the MEC server to a vehicle were also the same. Therefore, DAS-QF, where each of the uplink and downlink deadlines was set to the same value, i.e., 50 ms, worked well.

For the camera arrival ratio within the MC cycle deadline, DAS-QF could not obtain a high arrival ratio because it allocated the same uplink and downlink deadline value, i.e., 50 ms, even though the size of the chunk from the camera to the MEC server was larger than that of the chunk from the MEC server to the camera.

The reason that DAS-QF with DCF obtained a higher success ratio than DAS-QF was alone is that it could adaptively set each of the uplink and downlink deadlines, thus decreasing the number of chunk drops. Figures 4.5–4.8 show the delay distribution of chunks and its cumulative distribution function (CDF) for each case. The horizontal axis is the delay of each chunk, e.g., 30 ms on the horizontal axis means 20 ms $<$ delay $\leq$ 30 ms.

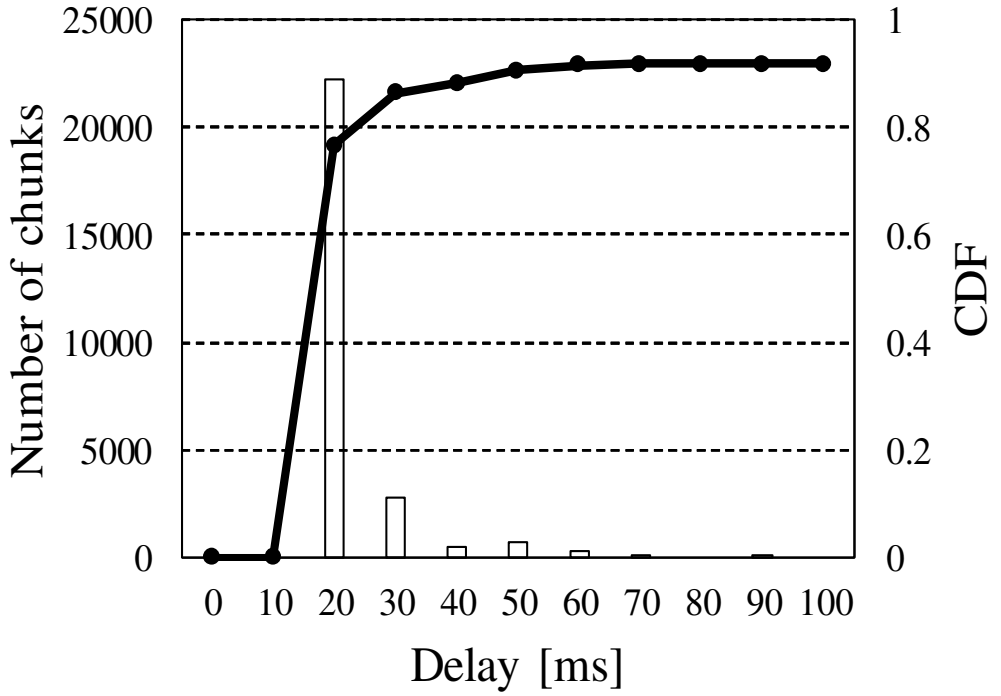

Figure 4.5: Delay distribution of DAS-QF alone (vehicle).

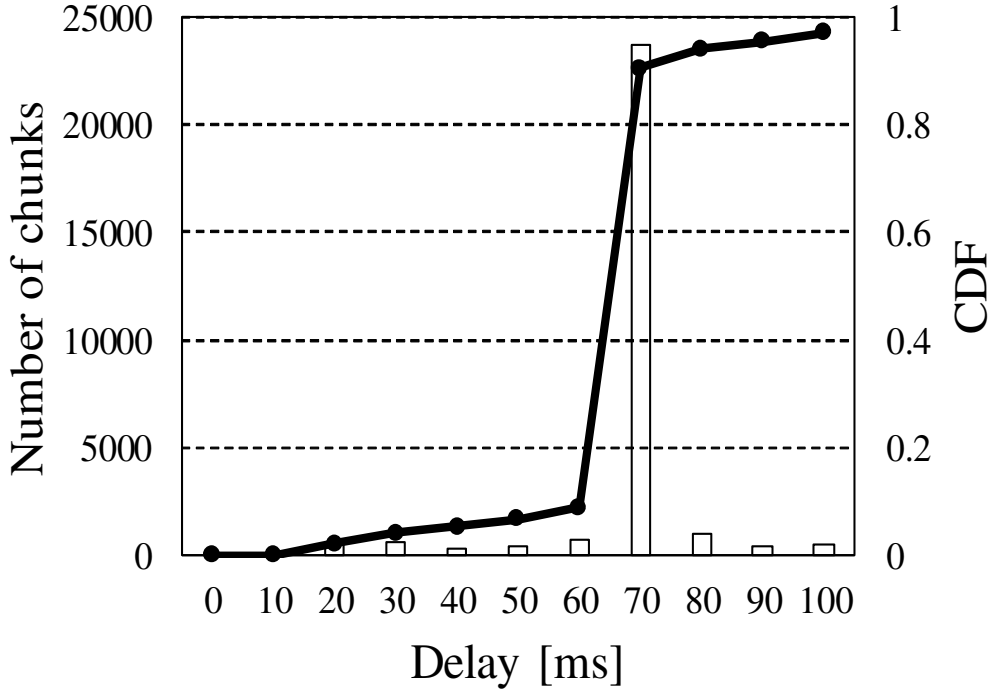

Figure 4.6: Delay distribution of DAS-QF with DCF (vehicle).

In the case of vehicles, the number of chunks that did not meet the MC cycle

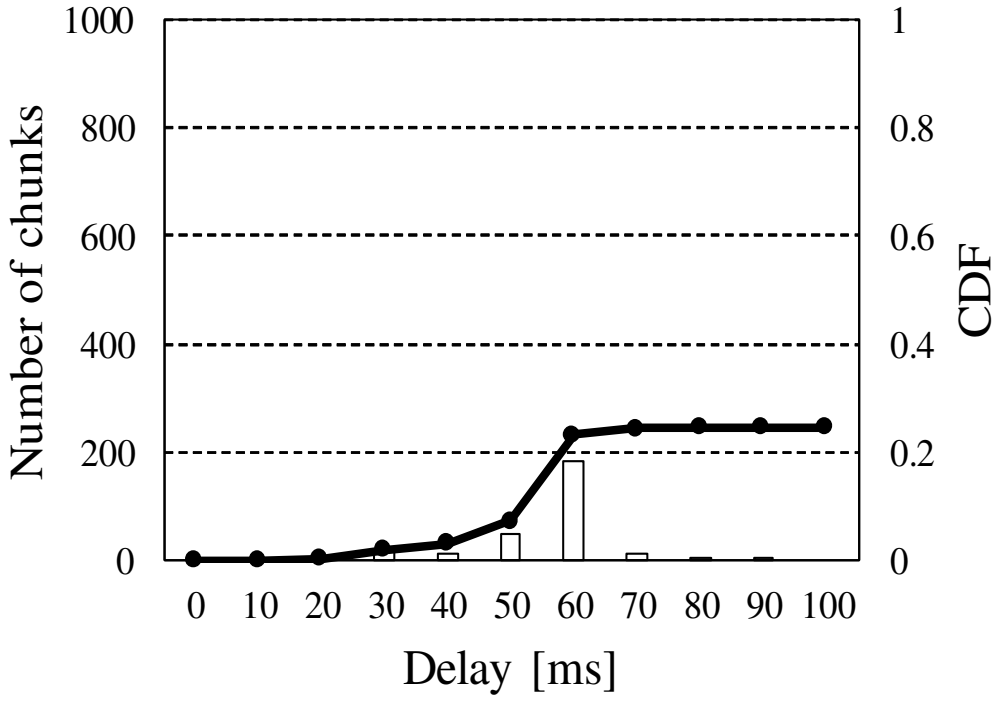

Figure 4.7: Delay distribution of DAS-QF alone (camera).

deadline for DAS-QF with DCF was approximately one-third that for DAS-QF alone, as shown by comparison of Figs. 4.5 and 4.6. In the case of cameras, the number of chunks that did not meet the MC cycle deadline for DAS-QF with DCF was approximately one-fifteenth that for DAS-QF alone, as shown by comparison of Figs. 4.7 and 4.8.

The peak delay distribution for vehicles existed at the delay of 20 ms, as shown in Fig. 4.5, when the DCF was not used. In contrast, when the DCF was used, the peak shifted from 20 ms to 70 ms, as shown in Fig. 4.6. This shift generated vacant radio resources that were then efficiently utilized by assigning them to camera traffic. DAS-QF with DCF shifted radio resources that are likely to be assigned to vehicles to cameras by setting the camera deadline to a smaller value. From Figs. 4.7 and 4.8, it can be confirmed that the vacant radio resources were assigned to camera traffic. Approximately 600 chunks did not meet the MC cycle deadline (100 ms) in Fig. 4.7.

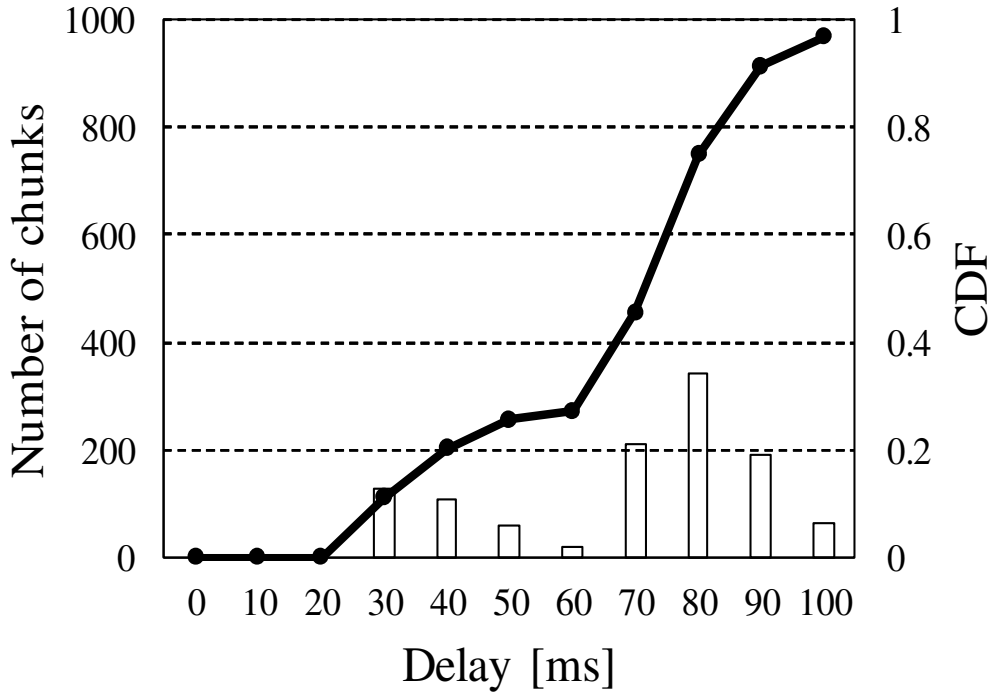

Figure 4.8: Delay distribution of DAS-QF with DCF (camera).

In Fig. 4.8, approximately 93% of these chunks were distributed to the left of 100 ms on the horizontal axis. These results demonstrate that DAS-QF with DCF could obtain higher arrival ratios within the MC cycle deadlines for both vehicles and cameras than DAS-QF alone.

The impact on BE cycle flow is now discussed. DAS-QF with DCF degraded the arrival rate for the BE cycle flow and average effective BE delay compared to the DAS-QF alone. The arrival rates for the BE cycle flow of DAS-QF with DCF and DAS-QF alone were 0.88 and 0.90, respectively. The average effective BE delays of DAS-QF with DCF and DAS-QF alone were 254 ms and 245 ms, respectively. This is because DAS-QF with DCF allocated less radio resources to the BE cycle flow than DAS-QF alone. In fact, the total bitrates received by all BE-UEs for DAS-QF with DCF and DAS-QF alone were 12.5 Mbps and 13.0 Mbps, respectively.

4.5 Summary

This chapter proposed a DCF that adaptively allocates uplink and downlink deadlines for MC cycle flows. The proposed DCF calculates the minimum coordination amount required to satisfy the arrival ratio within the MC cycle deadline and the uplink and downlink congestion levels. It updates the uplink and downlink deadlines based on their calculation results to avoid unintentional drops in an eNB.

This chapter investigated the DCF in a realistic environment by simulating an intersection where road traffic collisions often occur in ns-3 and then evaluating its performance. The results show that DAS-QF with DCF achieved a higher performance than DAS-QF alone. Specifically, the vehicle arrival ratio within the MC cycle deadline of DAS-QF with DCF achieved an improvement of approximately 5% compared to DAS-QF alone. In addition, the camera arrival ratio within the MC cycle deadline for DAS-QF with DCF achieved an improvement of approximately 70% compared to DAS-QF alone. The QoS requirements for connected car services were achieved through this improvement by applying the proposed DCF to DAS-QF.

Chapter 5

Deadline-Aware Bandwidth Assignment with MBR Control

5.1 Background and Outline

Mobile networks, such as LTE and 5G, are assumed to be used as platforms for real-time IoT applications [82, 83]. IoT devices, such as vehicles, cameras, drones, and sensors, exchange real-time information with each other to cooperate and work effectively with one another [84, 85].

In the IoT era, to satisfy the various QoS requirements on a shared mobile network with limited resources, adaptive control of network resources that considers each application behavior and communication environment is required because QoS requirements differ for each application [86]. Video streaming also focuses on QoE [87].

Model-based network resource management may miss deadlines due to the model uncertainties of mobile networks. For example, the throughput of an eNB is determined by a modulation scheme selected based on radio quality between each UE and the eNB. Different eNB vendors select different modulation schemes in accordance with radio quality; thus, mobile networks that include the eNBs of multiple vendors tend to have

disturbances in their throughput estimation for deadline-aware scheduling.

MBR control is one of the resource management methods for mobile networks. Ramamurthi et al. [62] proposed an MBR control method to enhance QoE in video streaming. This method conforms with 3GPP specifications and is useful for network resource management. However, there are currently no MBR-based resource management methods that support deadline-aware scheduling for mobile networks.

This dissertation proposes a deadline-aware resource management method based on MBR control with a feedback controller for delay-sensitive applications. The proposed method is suitable for practical use because MBR conforms with 3GPP specifications. The MBR can control both GBR and non-GBR traffic in the wireless uplink direction. GBR control is also a resource management method for mobile networks. However, GBR cannot control non-GBR traffic. The proposed deadline-aware resource management method uses MBR parameters to assume control of both GBR and non-GBR traffic. This research treats the model uncertainties as disturbances. The feedback controller maintains the effective throughput at a target level to meet the deadline even if disturbances exist. Numerical simulations demonstrate that the proposed method improves the success rate of data delivery within the deadline constraint in the presence of disturbances.

5.2 Mobile System and Its Modeling

The assumed mobile system is the LTE system shown in Fig. 5.1. Figure 5.2 shows the model of MBR control for an LTE uplink, and its parameters are listed in Table 5.1. This research assumes that MBR is used to control a mobile system. An MEC server calculates the MBR, and when the MEC server advertises the MBR to the PCRF, the

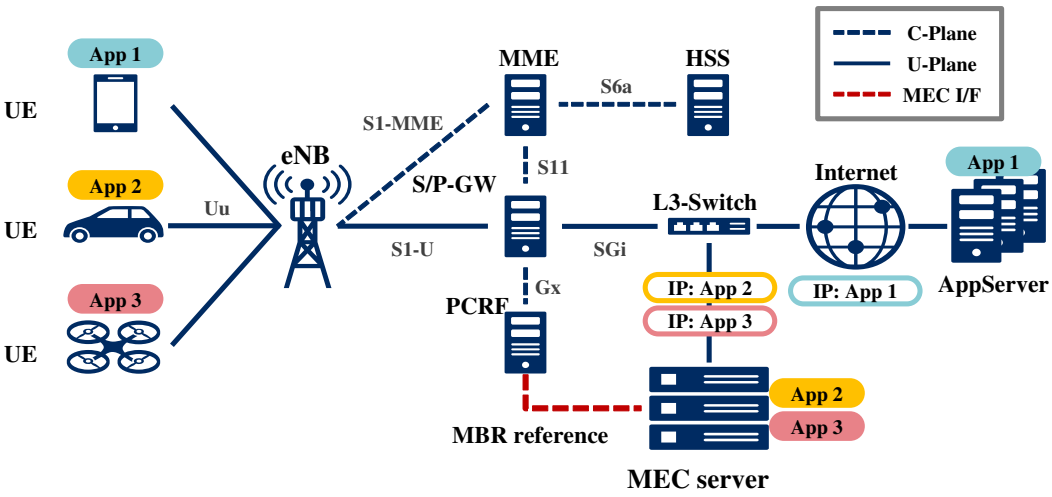

Figure 5.1: LTE system.

Table 5.1: System parameters

$r_i(k)$	Transmission rate of UE_i at sampling time k
$u_i(k)$	MBR reference allocated to UE_i at sampling time k
$d_i(k)$	Total disturbance in the dimension of throughput for UE_i at sampling time k
$\tau_i(k)$	Effective throughput of UE_i at sampling time k
$sd_{l,i}(k)$	Individual disturbance in the dimension of throughput for UE_i at sampling time k

PCRF sets the MBR to the mobile system. These functions are modeled as the MBR setting block in Fig. 5.2.

When UE_i sends a data chunk, its transmission rate at sampling time k, $r_i(k)$, is formulated as

$$r_i(k) = \begin{cases} u_i(k) & \text{(If } \mathrm{UE}_i \text{ has a chunk to be transmitted)} \\ 0 & \text{(Otherwise)} \end{cases}, \tag{5.1}$$

where $u_i(k)$ is the MBR reference allocated to UE_i at sampling time k. The effective throughput of UE_i at sampling time k, $\tau_i(k)$, might have a smaller value than the transmission rate of UE_i at sampling time k, $r_i(k)$, if the mobile system has some disturbances:

$$\tau_i(k) = r_i(k) - d_i(k) \tag{5.2}$$

.

A disturbance is considered to cause a difference between the target throughput and effective throughput. The mobile system might have various disturbances caused by the selection of different modulation schemes, fluctuation of radio quality, and other factors. Ideally, the mobile system should detect all disturbances and estimate their degrees. However, it is not realistic to consider all possible disturbances; thus, this research treats the summation of all disturbances as the total disturbance. It is advantageous that there is no need to estimate each disturbance separately, which results in a favorable implementation for mobile operators. The total disturbance in the dimension of throughput for UE_i at sampling time k, $d_i(k)$, is defined as

$$d_i(k) = sd_{1,i}(k) + sd_{2,i}(k) + sd_{3,i}(k), \tag{5.3}$$

where $sd_{1,i}(k)$, $sd_{2,i}(k)$, and $sd_{3,i}(k)$ denote disturbances caused by the selection of different modulation schemes, those caused by fluctuation of radio quality, and other disturbances, respectively.

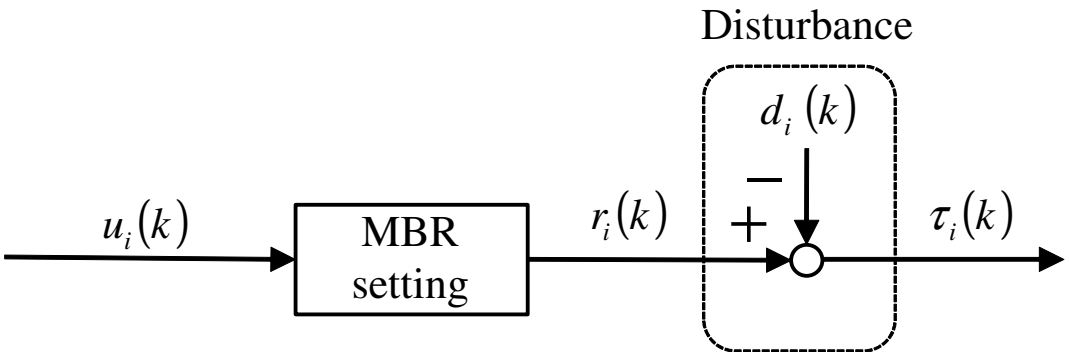

Figure 5.2: Model of MBR control for an LTE uplink.

Table 5.2: Parameters for deadline-aware MBR control.

$\tau_i^T(k)$	Target throughput of UE$_i$ at sampling time k
$r_{i,n}(k)$	Nominal throughput of UE$_i$ at sampling time k
K	Feedback gain

5.3 Deadline-Aware MBR Control

5.3.1 Static Controller

A block diagram of the MBR control system with a static controller is shown in Fig. 5.3, and additional parameters for deadline-aware MBR control are listed in Table 5.2. The static controller calculates the target throughput of UE$_i$ at sampling time k, $\tau_i^T(k)$, needed to meet the deadline. The target throughput of UE$_i$ at sampling time k, $\tau_i^T(k)$, is formulated as

$$\tau_i^T(k) = \frac{ChunkSize_i}{Deadline_i}, \tag{5.4}$$

where $ChunkSize_i$ and $Deadline_i$ are the chunk size of UE$_i$ and relative deadline from the time when a new chunk of UE$_i$ is generated, respectively. This research assumes that the deadline is shorter than the interval at which a chunk is generated.

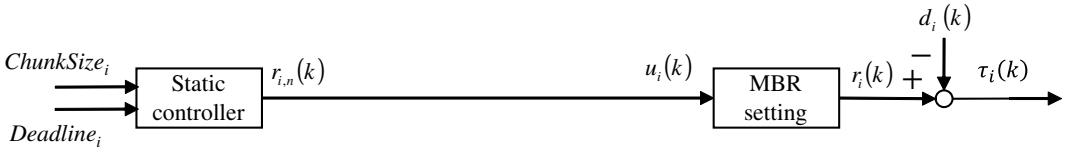

Figure 5.3: System with a static controller.

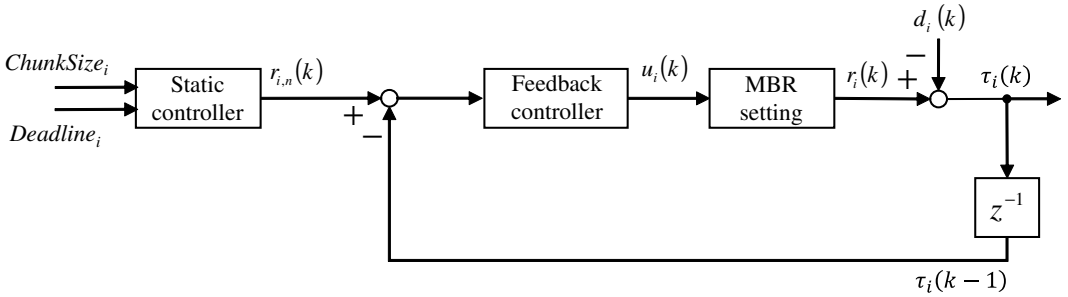

Figure 5.4: Proposed system with static and feedback controllers.

The nominal throughput of UE_i at sampling time k, $r_{i,n}(k)$, is defined as

$$r_{i,n}(k) = \begin{cases} \tau_i^T(k) & (t_i \leq Deadline_i) \\ 0 & (\text{Otherwise}) \end{cases}, \tag{5.5}$$

where t_i is the elapsed time from when a new chunk of UE_i was generated. This research sets $r_{i,n}(k)$ to $\tau_i^T(k)$ in the interval from when a chunk is generated to its deadline and sets $r_{i,n}(k)$ to zero in the interval from when the deadline of the chunk expires to when the next chunk is generated. When only a static controller is used in the mobile system, as shown in Fig. 5.3, the MBR reference of UE_i at sampling time k, $u_i(k)$, is equal to the output of the static controller, $r_{i,n}(k)$:

$$u_i(k) = r_{i,n}(k). \tag{5.6}$$

The mobile system cannot strictly allocate the target throughput because of the disturbance $d_i(k)$ within the system. The effective throughput is smaller than the target throughput, as shown in (5.2) and Fig. 5.3. Therefore, MBR control with only

a static controller cannot meet the deadline constraints.

5.3.2 Feedback Controller

A block diagram of the proposed MBR control system with static and feedback controllers is shown in Fig. 5.4. The MBR reference allocated to UE_i at sampling time k, $u_i(k)$, is formulated as

$$u_i(k) = K(r_{i,n}(k) - \tau_i(k-1)) + u_i(k-1), \tag{5.7}$$

where K is feedback gain.

The feedback controller adjusts the MBR reference $u_i(k)$ by comparing the target throughput $\tau_i^T(k)$ and effective throughput $\tau_i(k)$ to address the throughput degradation caused by the total disturbance $d_i(k)$. The proposed system has the advantage that there is no need to estimate disturbances $d_{1,i}(k)$, $d_{2,i}(k)$, and $d_{3,i}(k)$ separately, which results in a favorable implementation for mobile operators.

5.4 Numerical Simulation

In this section, the proposed resource control management is compared with resource control management without DC to evaluate the efficacy of DC.

5.4.1 Simulation Setup

This research compared a system with only a static controller with the proposed system with static and feedback controllers by numerical simulations in the presence of disturbances. The simulation topology is shown in Fig. 5.5. These simulations assumed there were three UEs connected to the same eNB: UE_1, which sent 150-Mb chunks at intervals of 3 s, with its deadline set to 1 s; UE_2, which sent 125-Mb

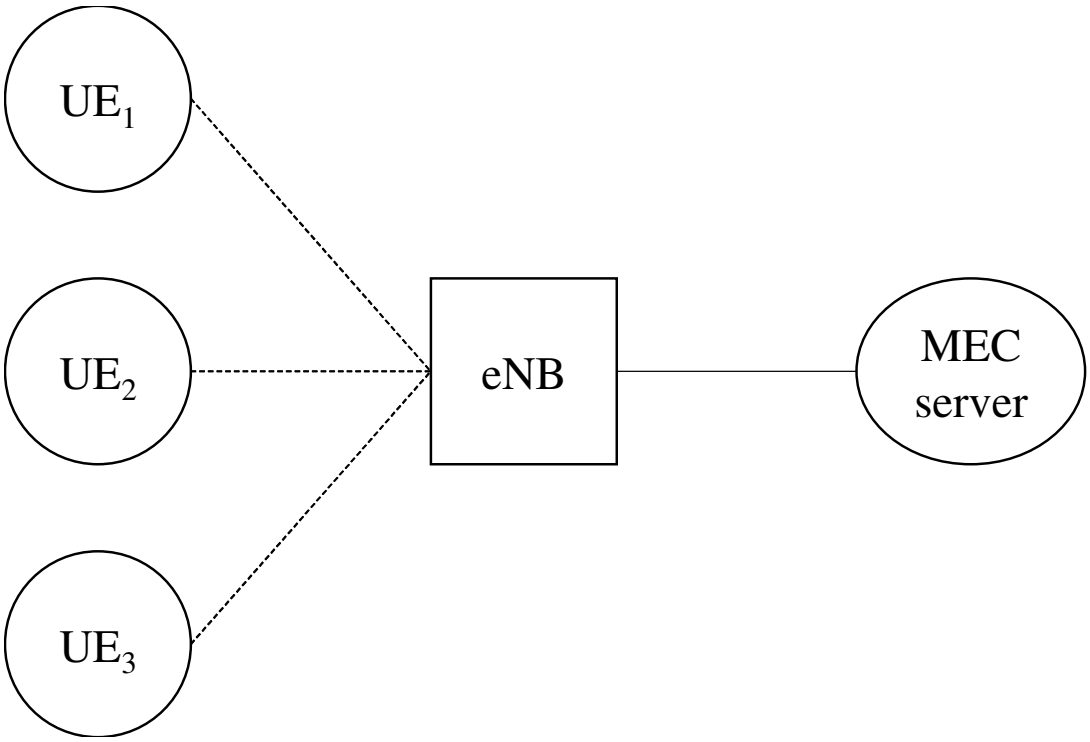

Figure 5.5: Simulation topology.

chunks at intervals of 4 s, with its deadline set to 2 s; and UE_3, which sent 145-Mb chunks at intervals of 5 s, with its deadline set to 3 s, as shown in Table 5.3. The simulation parameters are listed in Table 5.4. The three UEs shared an uplink with a bandwidth of 300 Mbps. The feedback gain K and control interval were set to 1 and 100 ms, respectively. It was assumed that an eNB performs radio resource allocation at intervals of 1 ms in accordance with bandwidth information from the proposed resource control management.

It was assumed that $d_{1,i}(k)$, $d_{2,i}(k)$, and $d_{3,i}(k)$ follow a normal distribution. Therefore, the total disturbance is formulated as

$$d_i(k) = (1 - NF(k))r_i(k), \tag{5.8}$$

where $NF(k)$ is the noise factor (NF) at sampling time k and follows a normal distribution. It is assumed that the range of $NF(k)$ is from 0 to 1.0. Therefore, when $NF(k)$ is calculated to be more than 1.0, it is set to 1.0. Similarly, when $NF(k)$ is

Table 5.3: UE settings.

UE	Period [s]	Chunk size [Mb]	Deadline [s]	Target throughput [Mbps]
UE_1	3	150	1	150
UE_2	4	125	2	62.5
UE_3	5	145	3	48.3

Table 5.4: Simulation parameters.

Bandwidth	300 Mbps
Control interval	0.1 s (1 step)
Number of steps	1000
K	1

calculated to be less than 0, it is set to 0.

5.4.2 Simulation Results

This section presents two sets of simulation results: throughput characteristics and success rate.

Throughput Characteristics

To evaluate the effectiveness of the feedback controller, the NF shown in Fig. 5.6 was applied to the simulation environment. For example, when time was 0–10 s, the NF is 0.8, as indicated by Fig. 5.6. This means that the effective throughput is the value

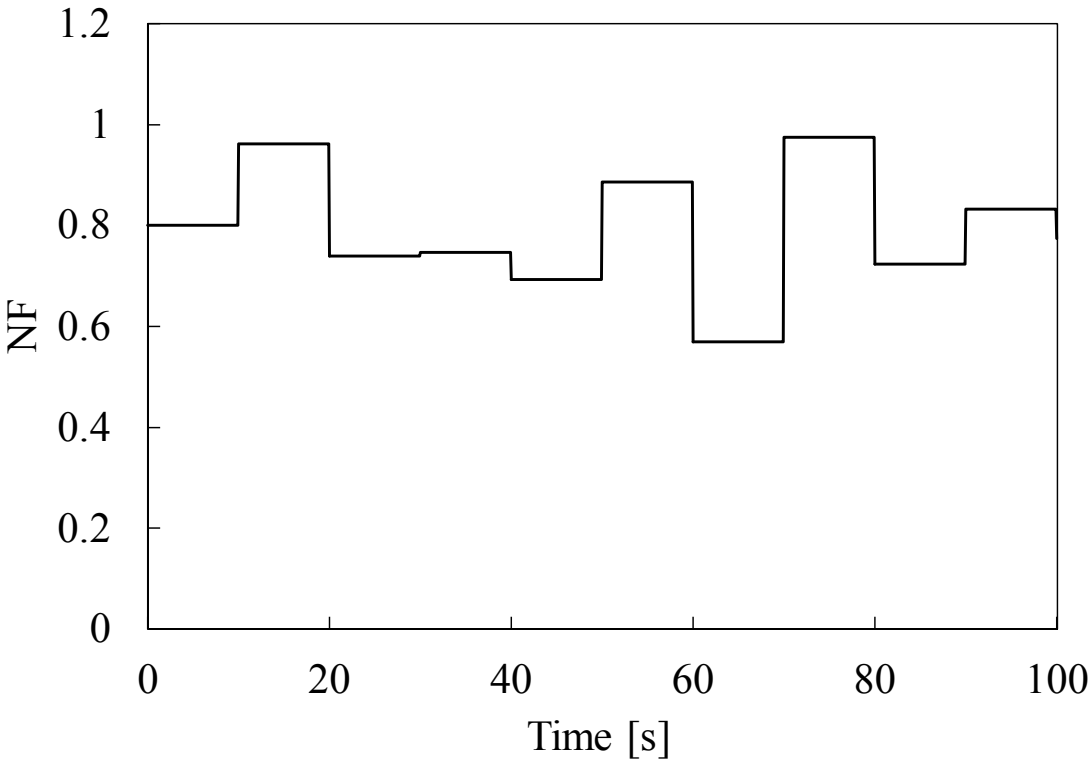

Figure 5.6: NF.

obtained by multiplying the NF and target throughput.

The throughput characteristics of each UE when the feedback controller is not used are shown in Figs. 5.7, 5.8, and 5.9. Data are delivered to their destinations within the deadline when the effective throughput is equal to the target throughput. As shown in Figs. 5.7, 5.8, and 5.9, when the feedback controller is not used, the effective throughput is lower than the target throughput. This means that a chunk generated by any UE cannot meet its deadline.

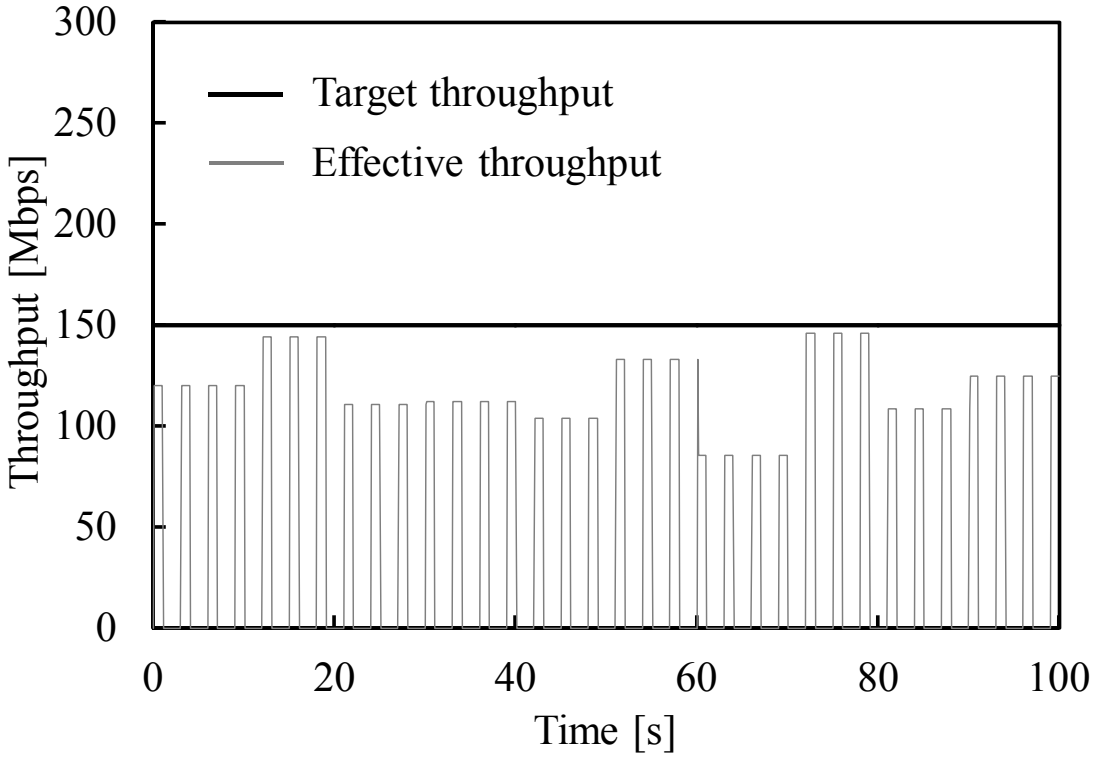

Figure 5.7: Effective throughput of UE_1 (without feedback controller).

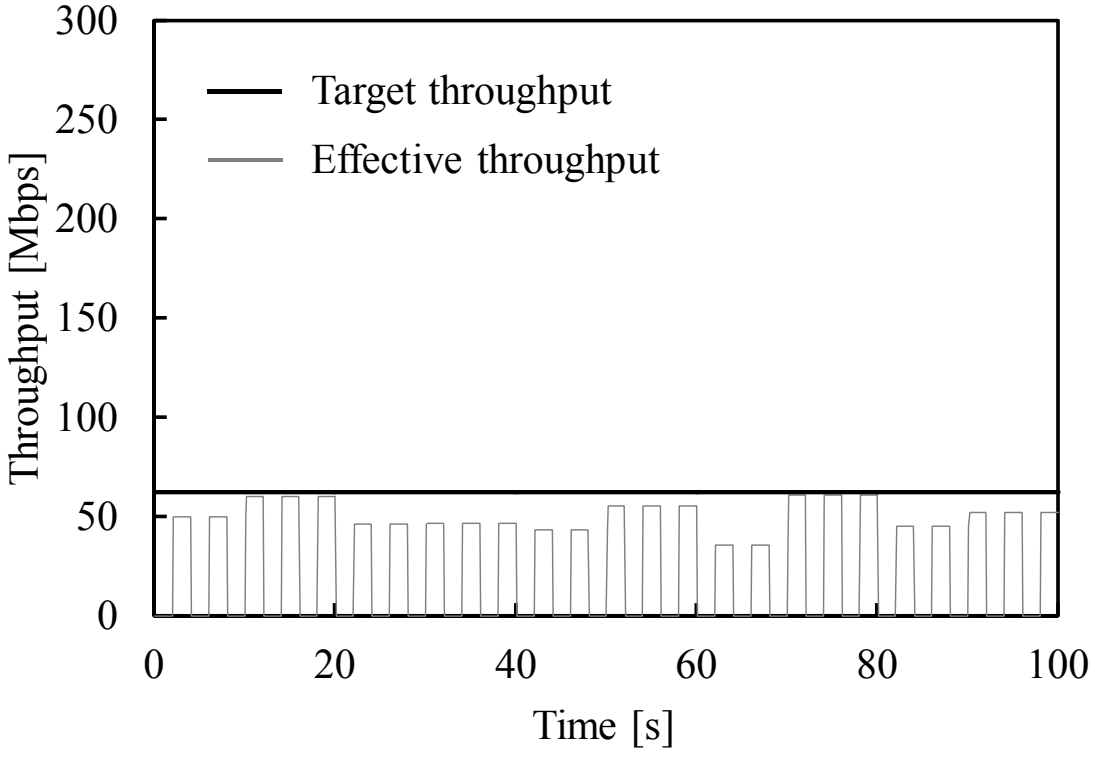

Figure 5.8: Effective throughput of UE_2 (without feedback controller).

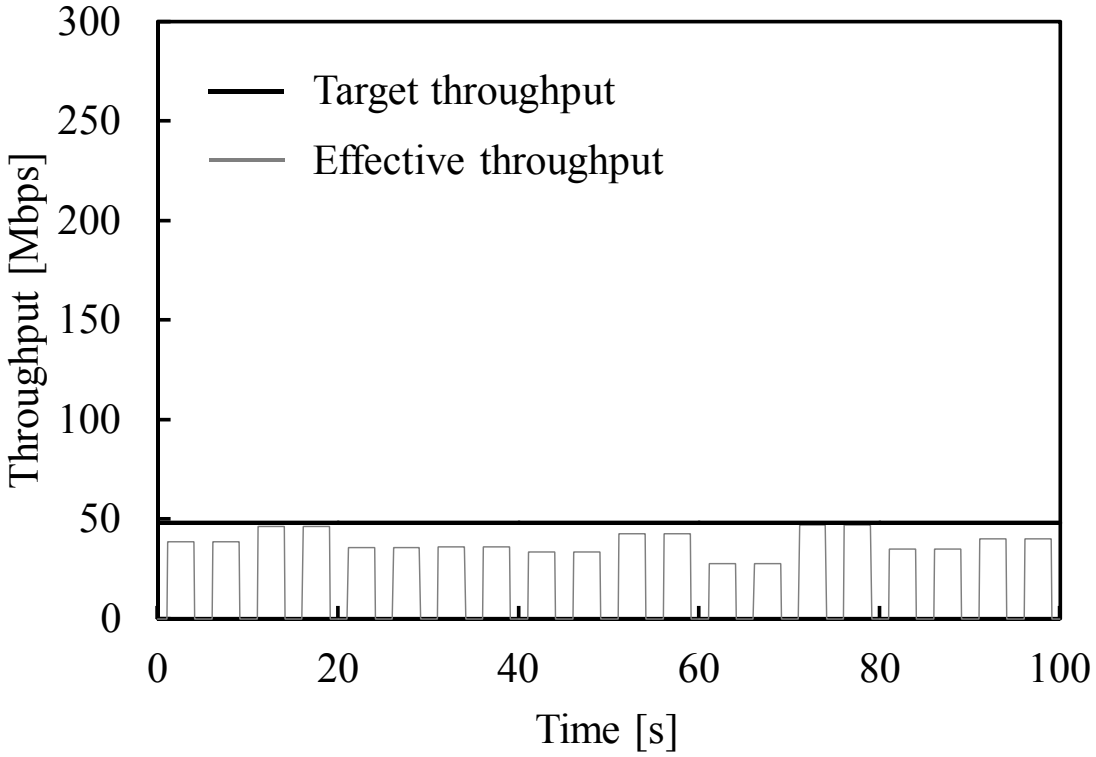

Figure 5.9: Effective throughput of UE_3 (without feedback controller).

The throughput characteristics of each UE when the feedback controller is used are shown in Figs. 5.10, 5.11, and 5.12.

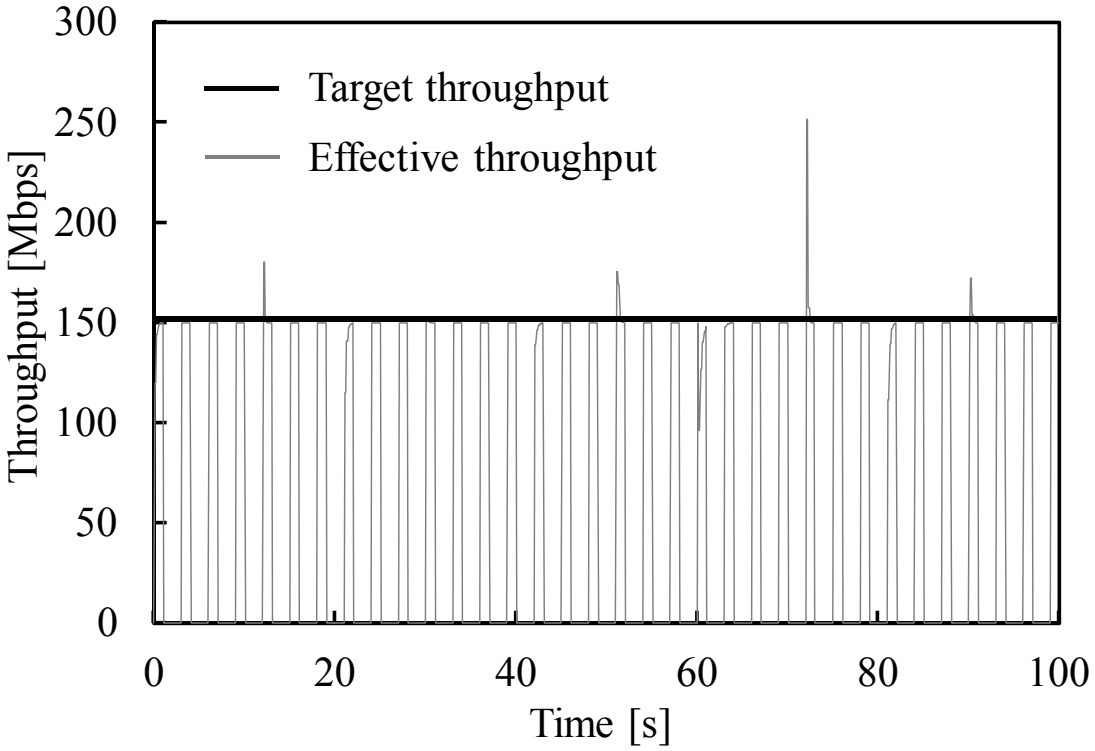

Figure 5.10: Effective throughput of UE_1 (with feedback controller).

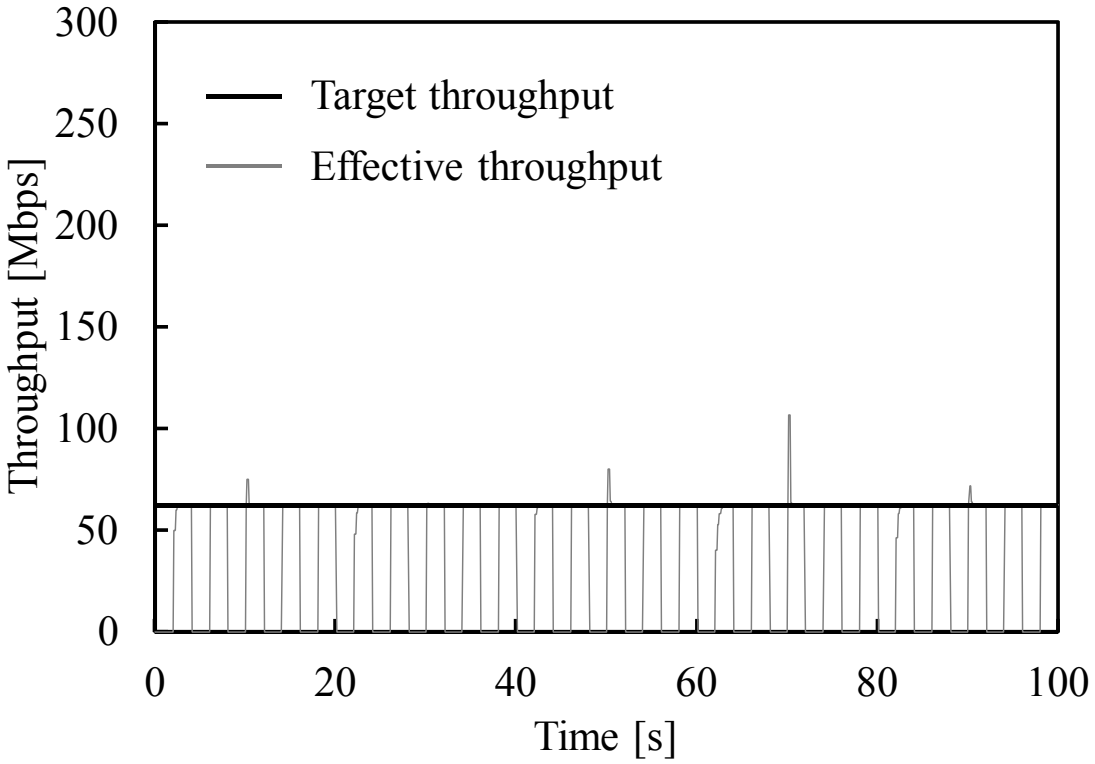

Figure 5.11: Effective throughput of UE_2 (with feedback controller).

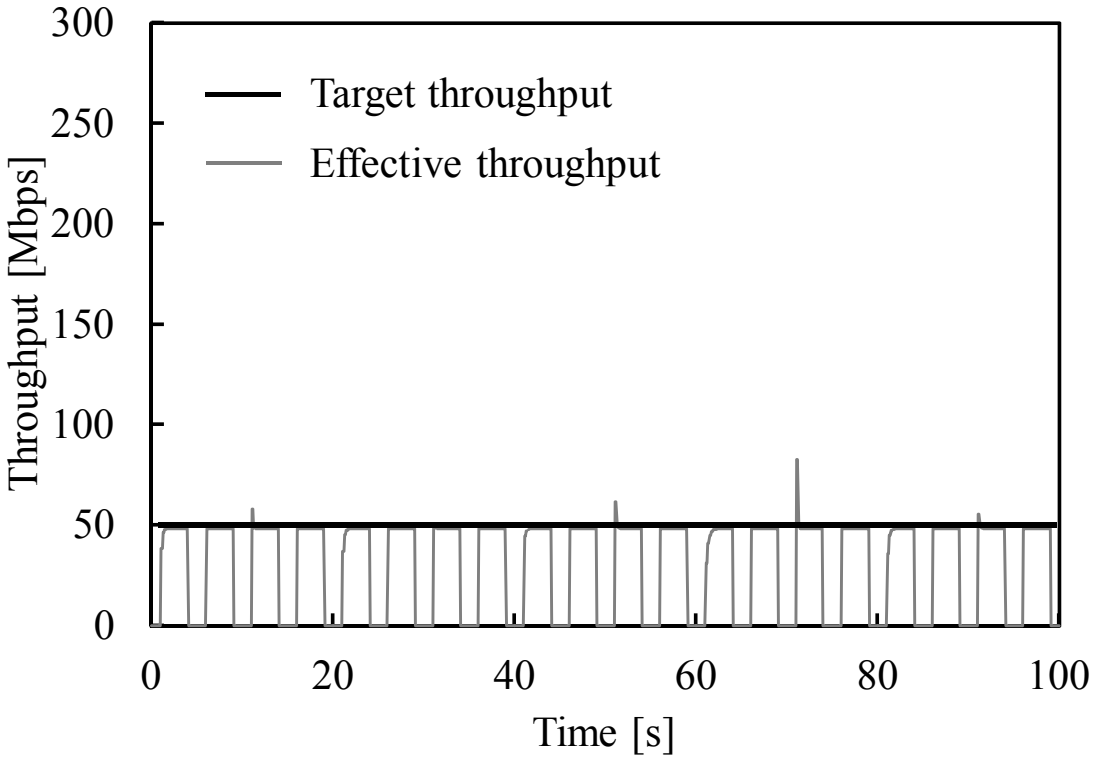

Figure 5.12: Effective throughput of UE_3 (with feedback controller).

The MBR reference is the set value of MBR compensated by the feedback controller. As shown in Figs. 5.10, 5.11, and 5.12, when the feedback controller is used, the effective throughput is closer to the target throughput compared with the proposed system without the feedback controller.

The proposed system with the feedback controller guarantees the deadline by adaptively adjusting the MBR reference that the proposed system with the feedback controller sets considering not only the deadline and chunk size but also disturbances. The effectiveness of the feedback controller is exemplified by the results for UE_1. As shown in Fig. 5.6, when time is 0–10 s, the NF is 0.8. When the feedback controller detects the noise level, it updates the MBR reference to 190 Mbps to achieve the target throughput. The feedback controller allocates the throughput required to meet its deadline even though the mobile system contains the NF, including disturbances.

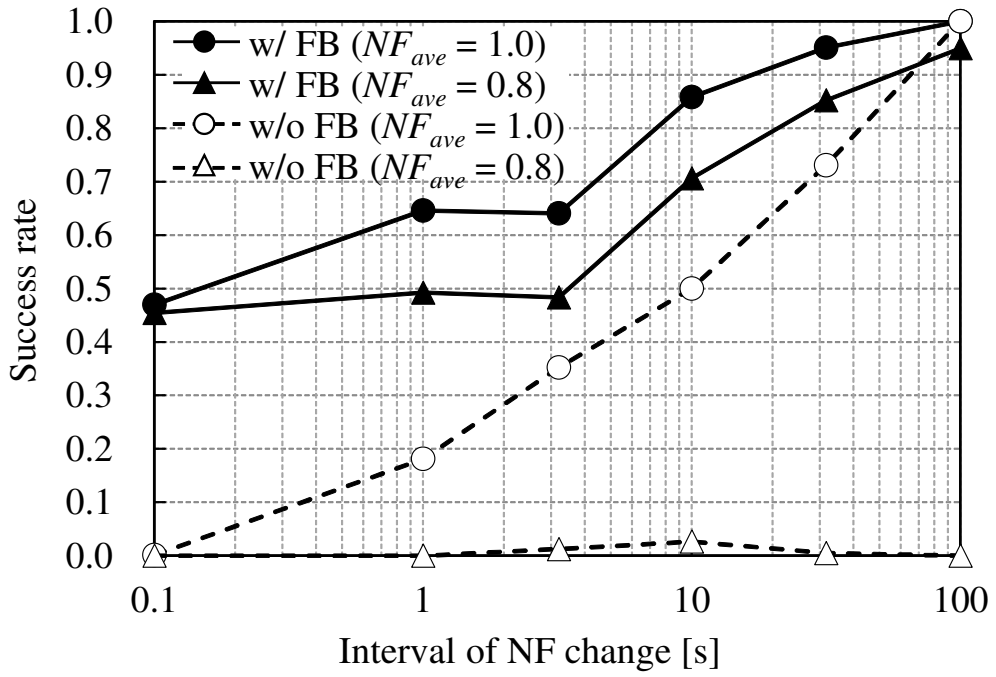

Figure 5.13: Success rate for each interval of NF change (standard deviation of NF = 0.1).

Figures 5.11 and 5.12 show that the compensation with the feedback controller achieves adaptive effective throughput for UE_2 and UE_3 as well as for UE_1.

Success Rate

The success rates are compared by changing the total disturbance to evaluate robustness to the total disturbance. The success rate $SuccessRate$ is formulated as

$$SuccessRate = \frac{\sum_i SuccessChunks_i}{\sum_i TotalChunks_i}, \tag{5.9}$$

where $SuccessChunks_i$ and $TotalChunk_i$ are the number of chunks meeting the deadline constraint for UE_i during the simulation time and the total number of chunks that UE_i sends during the simulation time, respectively.

The success rate for each interval of NF change is shown in Fig. 5.13. These are the simulation results of two cases where the average NF (NF_{ave} in the figure) is equal

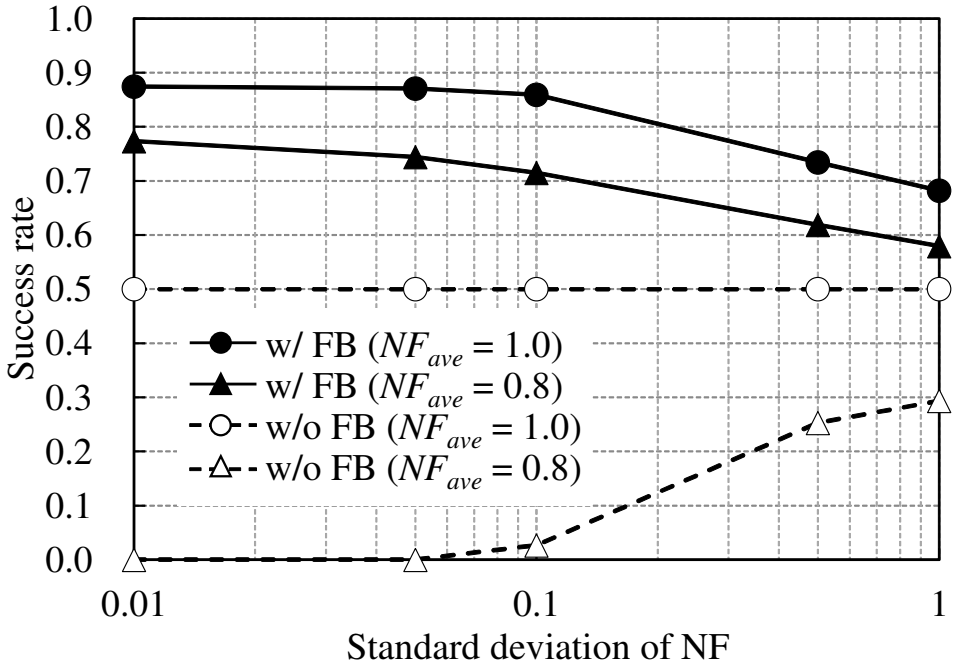

Figure 5.14: Success rate for standard deviation of NF (interval of NF change = 10 s).

to 0.8 and 1.0 when the standard deviation of NF is equal to 0.1. As can be seen, the proposed system with the feedback controller achieved a higher success rate than that without the feedback controller for both cases. This is because the feedback controller set the MBR reference by considering the throughput degradation caused by the total disturbance. However, even when the feedback controller was used, the shorter the interval of NF change was, the lower the success rate was. This is because the interval of NF change was shorter than the convergence time of the disturbance response.

The success rate for each standard deviation of NF is shown in Fig. 5.14. The interval of NF change is set to 10 s. As can be seen, the proposed system with the feedback controller achieved a higher success rate than that without the feedback controller for both cases. However, when the feedback controller was used, the larger the standard deviation of NF was, the lower the success rate was. This is because each MBR reference was set to a large value, and the summation of each MBR reference

exceeded the link bandwidth because the total disturbance was large. On the other hand, when the feedback controller was not used and the average of NF was equal to 0.8, the larger the standard deviation of NF was, the higher the success rate was. This is because, in this situation, the total disturbance was frequently equal to zero, as the standard deviation of NF was large.

5.5 Summary

This dissertation proposed a deadline-aware resource management method based on MBR control with a feedback controller to suppress the disturbances caused by model uncertainties of mobile networks. The results of numerical simulations revealed that the proposed method with the feedback controller provided a higher success rate than the proposed method without the feedback controller, even in the presence of disturbances. The proposed system has the advantage that there is no need to estimate each disturbance separately, which results in a favorable implementation for mobile operators. The proposed method is suitable for practical use because MBR conforms with 3GPP specifications. In the future, the proposed method should be evaluated to confirm its effectiveness in public cellular networks.

Chapter 6

Conclusions

This dissertation proposed DAS-QF, DAS-QF with DCF, and MBR control using a feedback controller for delay-sensitive applications to meet a certain deadline constraint.

In Chapter 3, DAS-QF, which decides priority in accordance with chunk size, cycle deadline, and radio quality, was proposed for delay-sensitive applications. With DAS-QF, chunks in a cycle flow were delivered within the cycle deadline. Simulation results showed that DAS-QF achieved higher performance than the PayDA and PF schedulers.

In Chapter 4, a DCF for DAS-QF was proposed to achieve higher performance than with DAS-QF alone. The DCF adaptively allocates uplink and downlink deadlines to each link in accordance with the cycle deadline and uplink and downlink congestion levels. The DCF is efficient for decreasing the number of chunks discarded by an eNB. Simulation results showed that DAS-QF with DCF achieved higher performance than DAS-QF alone.

Chapter 5 proposed a bandwidth assignment method with MBR control based on throughput feedback for delay-sensitive applications. This method treats the model

uncertainties as disturbances. The feedback controller in the proposed method maintains the effective throughput at a target level to meet the cycle deadline even if disturbances exist. Simulation results showed that MBR control with the feedback controller achieved higher performance than MBR control without the feedback controller. The proposed method has the advantage that there is no need to estimate each disturbance separately, which results in a favorable implementation for mobile operators. The proposed method is suitable for practical use because MBR conforms with 3GPP specifications.

This research contributes to the optimization of network resource management. Mobile networks, have evolved decade after decade, especially in terms of radio technologies. The development of radio technologies has enabled higher capacity, and this is expected to continue. 3GPP is still discussing sixth generation (6G) mobile networks to realize higher capacity and lower latency today. Mobile traffic data and the number of mobile users are assumed to continuously increase in the future. Therefore, the proposed network resource management methods, which efficiently allocate finite network resources, are important for realizing delay-sensitive applications, such as vehicle traffic collision avoidance and robotics automation, with mobile networks.

In this dissertation, network resource management methods for delay-sensitive applications were proposed under the LTE assumption, but this research is not limited to LTE. It is possible in principle that the proposed methods can be introduced to 5G or 6G networks. However, the implementation of the proposed methods on such networks has not been designed sufficiently. In the future, it should be considered how the proposed methods can be implemented in accordance with the standard specifications of 5G or 6G networks. The proposed methods must be evaluated to confirm their effectiveness in public cellular networks.

References

[1] Gabriel Brown, "Ultra-reliable low-latency 5G for industrial automation," *A Heavy Reading White Paper, Qualcomm*, 2018.

[2] Xichun Li, Abudulla Gani, Rosli Salleh and Omar Zakaria, "The future of mobile wireless communication networks," *Proceedings of the IEEE International Conference on Communication Software and Networks*, pp. 554–557, 2009.

[3] Hsiao-Hwa Chen, Mohsen Guizani and Werner Mohr, "Evolution toward 4G wireless networking [guest editorial]," *IEEE Network*, Vol. 21, No. 1, pp. 4–5, 2007.

[4] Stefan G. Hild, "A brief history of mobile telephony," *University of Cambridge Computer Laboratory Technical report*, UCAM-CL-TR-372, 1995.

[5] Verne H Mac Donald, "Advanced mobile phone service: The cellular concept," *The bell System Technical Journal*, Vol. 58, No. 1, pp. 15–41, 1979.

[6] Michel Mouly, Marie-Bernadette Pautet and Thomas Foreword By-Haug, "The GSM system for mobile communications," *Telecom Publishing*, 1992.

[7] Karl J Molnar and Gregory E Bottomley, "D-amps performance in pcs bands with array processing," *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, Vol. 3, pp. 1496–1500, 1996.

[8] John Capetanakis, "Generalized tdma: The multi-accessing tree protocol," *IEEE Transactions on Communications*, Vol. 27, No. 10, pp. 1476–1484, 1979.

[9] Jhong Sam Lee and Leonard E Miller, "CDMA systems engineering handbook," *Artech House, Inc.*, 1998.

[10] Shoichi Hirata, Akihisa Nakajima and Hisakazu Uesaka, "PDC mobile packet data communication network," *Proceedings of the 4th IEEE International Conference on Universal Personal Communications (ICUPC)*, pp. 644–648, 1995.

[11] Christian Bettstetter, Hans-Jorg Vogel and Jorg Eberspacher, "GSM phase 2+ general packet radio service GPRS: Architecture, protocols, and air interface," *IEEE Communications Surveys*, Vol. 2, No. 3, pp. 2–14, 1999.

[12] Erik Dahlman, Bjorn Gudmundson, Mats Nilsson and A Skold, "UMTS/IMT-2000 based on wideband cdma," *IEEE Communications Magazine*, Vol. 36, No. 9, pp. 70–80, 1998.

[13] Harri Holma and Antti Toskala, "WCDMA for UMTS: Radio access for third generation mobile communications," *John Wiley & Sons*, 2002.

[14] Vijay K Garg, "IS-95 CDMA and CDMA2000: Cellular/PCS systems implementation," *Pearson Education*, 1999.

[15] 3GPP2 C.S0029, "cdma2000 high rate packet data air interface specifcation," 2002.

[16] Erik Dahlman, Stefan Parkvall, Johan Skold and Per Beming, "3G evolution: HSPA and LTE for mobile broadband," *Academic Press*, 2010.

[17] Stefania Sesia, Issam Toufik and Matthew Baker, “LTE-the UMTS long term evolution: from theory to practice,” *John Wiley & Sons*, 2011.

[18] 3GPP TS36.300, “Evolved universal terrestrial radio access (e-utra) and evolved universal terrestrial radio access network (e-utran); overall description,” v8.12.0, 2010.

[19] Amitava Ghosh, Rapeepat Ratasuk, Bishwarup Mondal, Nitin Mangalvedhe and Tim Thomas, “Lte-advanced: next-generation wireless broadband technology,” *IEEE Wireless Communications*, Vol. 17, No. 3, pp. 10–22, 2010.

[20] Shao-Yu Lien, Shin-Lin Shieh, Yenming Huang, Borching Su, Yung-Lin Hsu and Hung-Yu Wei, “5G new radio: Waveform, frame structure, multiple access, and initial access,” *IEEE Communications Magazine*, Vol. 55, No. 6, pp. 64–71, 2017.

[21] Hyoungju Ji, Sunho Park, Jeongho Yeo, Younsun Kim, Juho Lee and Byonghyo Shim, “Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects,” *IEEE Wireless Communications*, Vol. 25, No. 3, pp. 124–130, 2018.

[22] Chih-Ping Li, Jing Jiang, Wanshi Chen, Tingfang Ji and John Smee, “5G ultra-reliable and low-latency systems design," *Proceedings of the IEEE European Conference on Networks and Communications (EuCNC)*, pp. 1–5, 2017.

[23] “3GPP specification series: 38 series,” available from < https://www.3gpp.org/dynareport/38-series.htm >, (accessed 2020-07-18).

[24] 3GPP TR 38.101-1, “NR; user equipment (UE) radio transmission and reception; part 1: Range 1 standalone,” 2020.

[25] 3GPP TR 38.101-2, “NR; user equipment (ue) radio transmission and reception; part 2: Range 2 standalone,” 2020.

[26] Paul Nikolich, C Lin, Jouni Korhonen, Roger Marks, Blake Tye, Gang Li, Jiqing Ni and Siming Zhang, “Standards for 5G and beyond: Their use cases and applications,” *IEEE 5G Tech Focus*, Vol. 1, No. 2, 2017.

[27] Gerhard Fettweis and Siavash Alamouti, “ 5G: Personal mobile internet beyond what cellular did to telephony,” *IEEE Communications Magazine*, Vol. 52, No. 2, pp. 140–145, 2014.

[28] ITU-T, “The tactile internet: ITU-T technology watach report,” *Technical report*, 2014.

[29] 3GPP TS22.804, “Study on communication for automation in vertical domains,” v0.3.0, 2017.

[30] 3GPP R1-1609664, “Comparison of slot and mini-slot based approaches for URLLC, *Technical report*, 2016.

[31] Zexian Li, Mikko A Uusitalo, Hamidreza Shariatmadari and Bikramjit Singh, “5G urllc: Design challenges and system concepts,” *Proceeding of the IEEE 15th International Symposium on Wireless Communication Systems (ISWCS)*, pp. 1–6, 2018.

[32] World Health Organization, “Global status report on road safety 2015”, 2015.

[33] “ITARDA” (online), available from < http://www.itarda.or.jp/english/ >, (accessed 2020-07-18).

[34] Mengfei Xie, Yong Shang, Zhenyu Yang, Yi Jing and Haijun Zhou, "A novel mbsfn scheme for vehicle-to-vehicle safety communication based on LTE network," *Proceedings of the IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, pp. 1–5, 2015.

[35] Shanzhi Chen, Jinling Hu, Yan Shi and Li Zhao, "LTE-V: A TD-LTE-based V2X solution for future vehicular network, *IEEE Internet of Things journal*, Vol. 3, No. 6, pp. 997–1005, 2016.

[36] Shao-hui Sun, Jin-ling Hu, Ying Peng, Xue-ming Pan, Li Zhao and Jia-yi Fang, "Support for vehicle-to-everything services based on LTE," *IEEE Wireless Communications*, Vol. 23, No. 3, pp. 4–8, 2016.

[37] S Vishnu, Ullas Ramanadhan, Nirmala Vasudevan and Anand Ramachandran, "Vehicular collision avoidance using video processing and vehicle-to-infrastructure communication," *Proceedings of the IEEE International Conference on Connected Vehicles and Expo (ICCVE)*, pp. 387–388, 2015.

[38] James Temperton, "One nation under cctv: the future of automated surveillance" (online), available from < https://www.wired.co.uk/article/one-nation-under-cctv >, (accessed 2020-07-18).

[39] Frank Langfitt, "In china, beware: A camera may be watching you" (online), available from < https://www.npr.org/2013/01/29/170469038/in-china-beware-a-camera-may-be-watching-you>, (accessed 2020-07-18).

[40] 3GPP TR 22.885, "Study on LTE support for vehicle to everything (V2X) services," *Technical Report*, 2015.

[41] Nobuhiko Itoh, Motoki Morita, Takanori Iwai, Kozo Satoda and Ryogo Kubo, "A deadline-aware scheduling scheme for connected car services using mobile networks with quality fluctuation," *IEICE Transactions on Communications*, Vol. E102.B, No. 3, pp. 474–483, 2019.

[42] Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Congestion-adaptive and deadline-aware scheduling for connected car services over mobile networks," *IEICE Transactions on Communications*. (advance publication)

[43] Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Maximum bit rate control based on throughput feedback for deadline-aware resource management in cellular networks," *IEICE Communications Express*, Vol. 9, No. 7, pp. 250–255, 2020.

[44] 3GPP TR 23.401, "General packet radio service (gprs) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access," *Technical Report*, 2019.

[45] 3GPP TR 23.203, "Policy and charging control architecture," *Technical Report*, 2019.

[46] 3GPP TR 32.240, "Telecommunication management; charging management; charging architecture and principles," *Technical Report*, 2020.

[47] 3GPP TR 32.296, "Telecommunication management; charging management; online charging system (OCS): Applications and interfaces," *Technical Report*, 2018.

[48] Raymond Kwan, Cyril Leung and Jie Zhang, "Proportional fair multiuser scheduling in LTE," *IEEE Signal Processing Letters*, Vol. 16, No. 6, pp. 461–464, 2009.

[49] S.-B. Lee, I. Pefkianakis, A. Meyerson, S. Xu and S. Lu, "Proportional fair frequency-domain packet scheduling for 3gpp LTE uplink," *Proceedings of the IEEE INFOCOM*, pp. 2611-2615, 2009.

[50] Raj Jain, Arjan Durresi and Gojko Babic, "Throughput fairness index: An explanation," *ATM Forum Contribution*, Vol. 99-0045, 1999.

[51] Erik Dahlman, Stefan Parkvall, Johan Skold and Per Beming, "3G Evolution 2nd Edition," *Elsevier Science Publishing Company Incorporated*, 2008.

[52] Li-Chun Wang and Wei-Jun Lin, "Throughput and fairness enhancement for ofdma broadband wireless access systems using the maximum C/I scheduling," *Proceedings of the IEEE 60th Vehicular Technology Conference (VTC2004-Fall)*, Vol. 7, pp. 4696–4700, 2004.

[53] Yifan Zhang, Muqing Wu, Rui Zhang, Panfeng Zhou and Shiping Di, "Adaptive QoS-aware resource allocation for high-speed mobile LTE wireless systems," *Proceedings of the IEEE Global Communications Conference (Globecom) Workshops*, pp.947–952, 2013.

[54] Giuseppe Piro, Luigi Alfredo Grieco, Gennaro Boggia, Rossella Fortuna and Pietro Camarda, "Two-level downlink scheduling for real-time multimedia services in LTE networks," *IEEE Transactions on Multimedia*, Vol. 13, No. 5, pp. 1052–1065, 2011.

[55] C. L. Liu and James W. Layland, "Scheduling algorithms for multiprogramming in a hard-real-time environment," *Journal of the ACM*, Vol. 20, No. 1, pp. 46–61, 1973.

[56] Matteo Maria Andreozzi, Giovanni Stea, Andrea Bacioccola and Roberto Rossi,"Flexible scheduling for real-time services in high-speed packet access cellular networks," *Proceedings of the IEEE European Wireless Conference*, 2009.

[57] Khaled MF Elsayed and Ahmed KF Khattab, "Channel-aware earliest deadline due fair scheduling for wireless multimedia networks," *Wireless Personal Communications*, Vol. 38, No. 2, pp. 233–252, 2006.

[58] Bin Liu, Hui Tian and Lingling Xu, "An efficient downlink packet scheduling algorithm for real time traffics in LTE systems," *Proceeding of the IEEE 10th Consumer Communications and Networking Conference (CCNC)*, pp. 364–369, 2013.

[59] Nusrat Afrin, Jason Brown and Jamil Y Khan, "A packet age based LTE uplink packet scheduler for m2m traffic," *Proceedings of the IEEE 7th International Conference on Signal Processing and Communication Systems (ICSPCS)*, pp. 1–8, 2013.

[60] Marcus Haferkamp, Benjamin Sliwa, Christoph Ide and Christian Wietfeld, "Payload-size and deadline-aware scheduling for time-critical cyber physical systems," *Proceedings of the IEEE 2017 Wireless Days*, pp. 4–7, 2017.

[61] 3GPP TS 29.212, "Policy and charging control (pcc) over gx reference point," *Technical report*, 2015.

[62] Vishwanath Ramamurthi, Ozgur Oyman and Jeffrey Foerster, "Video-qoe aware resource management at network core," *Proceedings of the IEEE Global Communications Conference, GLOBECOM'14*, pp. 1418–1423, 2014.

[63] “SONPO” (online), available from < http://www.sonpo.or.jp/efforts/reduction/ >, (accessed 2016-03-10). (in Japanese)

[64] Takanori Iwai, Daichi Kominami, Masayuki Murata, Ryogo Kubo and Kozo Satoda, “Mobile network architectures and context-aware network control technology in the IoT era,” *IEICE Transactions on Communications*, Vol. E101.B, No. 10, pp. 2083–2093, 2018.

[65] Ke Zhang, Yuming Mao, Supeng Leng, Quanxin Zhao, Longjiang Li, Xin Peng, Li Pan, Sabita Maharjan and Yan Zhang, “Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks,” *IEEE access*, Vol. 4, pp. 5896–5907, 2016.

[66] Tarik Taleb, Konstantinos Samdanis, Badr Mada, Hannu Flinck, Sunny Dutta and Dario Sabella, “On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration,” *IEEE Communications Surveys & Tutorials*, Vol. 19, No. 3, pp. 1657–1681, 2017.

[67] Pawani Porambage, Jude Okwuibe, Madhusanka Liyanage, Mika Ylianttila and Tarik Taleb, “Survey on multi-access edge computing for internet of things realization,” *IEEE Communications Surveys & Tutorials*, Vol. 20, No. 4, pp. 2961–2991, 2018.

[68] Yuan Ai, Mugen Peng and Kecheng Zhang, “Edge computing technologies for internet of things: a primer,” *Elsevier Digital Communications and Networks*, Vol. 4, No. 2, pp. 77–86, 2018.

[69] Cisco Visual Networking, “Cisco visual networking index: Global mobile data traffic forecast update 2017-2022,” *Cisco White Paper*, 2016.

[70] Ashkan Nikravesh, David R Choffnes, Ethan Katz-Bassett, Z Morley Mao and Matt Welsh, "Mobile network performance from user devices: A longitudinal, multidimensional analysis," *Proceedings of International Conference on Passive and Active Network Measurement, Springer*, pp. 12–22, 2014.

[71] Saba Siraj, Ajay Kumar Gupta and Rinku-Badgujar, "Network simulation tools survey," *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 1, No. 4, pp. 201–210, 2012.

[72] Michael Behrisch, Laura Bieker, Jakob Erdmann and Daniel Krajzewicz, "SUMO—simulation of urban mobility: An overview," *Proceedings of the Third International Conference on Advances in System Simulation*, Vol. 2011, 2011.

[73] Michael Behrisch, Daniel Krajzewicz and Melanie Weber, "Simulation of Urban Mobility," *Springer Berlin Heidelberg*, 2014.

[74] Giuseppe Piro, Nicola Baldo and Marco Miozzo, "An LTE module for the ns-3 network simulator," *Proceeding of the 4th International ICST Conference on Simulation Tools and Techniques, SIMUTools '11,* pp. 415–422, 2011.

[75] 3GPP TS 24.386, "User equipment (UE) to V2X control function," *Technical Report*, 2017.

[76] Ahmad M El-Hajj and Zaher Dawy, "On delay-aware joint uplink/downlink resource allocation in ofdma networks," *Proceeding of the IEEE 2013 Symposium on Computers and Communications (ISCC)*, pp. 000257–000262, 2013.

[77] Ahmad M El-Hajj, Zaher Dawy and Walid Saad, "A stable matching game for joint uplink/downlink resource allocation in OFDMA wireless networks,"

Proceeding of the IEEE International Conference on Communications (ICC), pp. 5354–5359, 2012.

[78] 3GPP TS 22.185, "Service requirements for v2x services," *Technical Report*, 2017.

[79] Ravi Teja Mullapudi, Steven Chen, Keyi Zhang, Deva Ramanan and Kayvon Fatahalian, "Online model distillation for efficient video inference," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3573–3582, 2019.

[80] Joseph Redmon and Ali Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[81] 3GPP TS 36.213, "Physical layer procedures (release 10)," *Technical Report*, 2011.

[82] Angel Molina-Garcia, Juan Alvaro Fuentes, Emilio Gomez-Lazaro, Alberto Bonastre, José Carlos Campelo and Juan José Serrano, "Development and assessment of a wireless sensor and actuator network for heating and cooling loads," *IEEE Transactions on Smart Grid*, Vol. 3, No. 3, pp. 1192–1202, 2012.

[83] Lu Tan and Neng Wang, "Future internet: The internet of things," *Proceedings of the IEEE 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, pp. V5-376–V5-380, 2010.

[84] Lukas Marcel Schalk and Martin Herrmann, "Suitability of lte for drone-to-infrastructure communications in very low level airspace," *Proceedings of the IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, pp. 1–7, 2017.

[85] Yan Hu, Jingjing Feng and Wenli Chen, "A LTE-cellular-based V2X solution to future vehicular network," *Proceedings of the IEEE 2nd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pp. 2658–2662, 2018.

[86] Ilias Gravalos, Prodromos Makris and Konstantinos Christodoulopoulos, "Efficient network planning for internet of things with QoS constraints," *IEEE Internet of Things Journal*, Vol. 5, No. 5, pp. 3823-3836, 2018.

[87] Xiaoqi Yin, Abhishek Jindal, Vyas Sekar and Bruno Sinopoli, "A control-theoretic approach for dynamic adaptive video streaming over http," *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, pp. 325–338, 2015.

List of Achievements

Journals (First Author)

1. Nobuhiko Itoh, Motoki Morita, Takanori Iwai, Kozo Satoda and Ryogo Kubo, "A deadline-aware scheduling scheme for connected car services using mobile networks with quality fluctuation," *IEICE Transactions on Communications*, Vol. E102-B, No. 3, pp. 474–483, March 2019.

2. Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Maximum bit rate control based on throughput feedback for deadline-aware resource management in cellular networks," *IEICE Communications Express*, Vol. 9, No. 7, pp. 250–255, July 2020.

3. Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Congestion-adaptive and deadline-aware scheduling for connected car services over mobile networks," *IEICE Transactions on Communications*. (advance publication)

Journals (Co-Author)

1. Yuta Yamamoto, Nobuhiko Itoh and Miki Yamamoto, "Performance improvement of proxy-based TCP in tandem ad hoc networks," *IEICE Transactions on Communications*, Vol. J93-B, No. 2, pp. 191–200, February 2010.

International Conferences

1. Nobuhiko Itoh and Miki Yamamoto, "Proxy-based TCP with adaptive rate control and intentional flow control in ad hoc networks," *Proceedings of the IEEE Global Communications Conference, GLOBECOM '08, New Orleans, LA, USA*, pp. 1-6, November 2008.

2. Xiaoyuan Cao, Noboru Yoshikane, Takehiro Tsuritani, Sota Yoshida, Shan Gao, Masaki Tanaka, Ryo Mikami, Takaya Miyazawa, Nobuhiko Itoh, Yoshihiro Isaji, Takehiro Sato, Satoru Okamoto, Hideo Iwamoto, Toru Katagiri, Takamasa Kobayashi, Naoki Kakegawa, Tsutomu Kawada, Shinji Tanaka, Rie Hayashi, Takafumi Hamano and Kouichi Genda, "Nationwide demonstration of software defined optical transport networking via multi-domain orchestration," *Proceedings of 18th Annual International Conference Next Generation Internet and Related Technologies (SDN/MPLS)*, November 2015.

3. Nobuhiko Itoh, "RISE 3.0: The design and implementation of SDN/OpenFlow testbed considering node capacity and inflexible topology," *APAN 38th Future Internet Testbed Working Group, Nantou, Taiwan,* August 2014.

4. Nobuhiko Itoh, "RISE - SDN testbed in JGN-X," *Internet2 Technology Exchange, Indianapolis, USA,* October 2014.

5. Nobuhiko Itoh, "RISE - SDN testbed on JGN-X," *APII Workshop, Osaka, Japan,* November 2014.

6. Nobuhiko Itoh, "RISE wide-area SDN testbed," *Roadmap to Operating SDN-based Networks Workshop, California,* July 2015.

7. Nobuhiko Itoh, "Next RISE architecture," *APAN 40th Future Internet Testbed Working Group, Malaysia,* August 2015.

8. Nobuhiko Itoh, "Introduction of high scalable and flexible RISE testbed with ODENOS and Lagopus," *APAN 41th Future Internet Testbed Working Group, Manila, Fhilippines,* January 2016.

9. Nobuhiko Itoh, Hiroya Kaneko, Akihito Kohiga, Takanori Iwai and Hideyuki Shimonishi, "Novel packet scheduling for supporting various real-time IoT applications for LTE networks," *Proceedings of the IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR 2017), Naples, FL, USA*, May 2017.

Domestic Conferences

1. Nobuhiko Itoh and Miki Yamamoto, " A study on cache placement method considering mobility in ad hoc networks," *IEICE Kansai Technical Report, B1-3,* March 2006. (in Japanese)

2. Nobuhiko Itoh, Yuki Imanishi and Miki Yamamoto, "Performance evaluation of TCP proxy in ad hoc networks," *Proceedings of the 2007 IEICE Society Conference,* Vol. 2 p. 54, September 2007. (in Japanese)

3. Nobuhiko Itoh, Yuki Imanishi and Miki Yamamoto, "Performance evaluation of TCP proxy in ad hoc networks," *IEICE Technical Report, Network System,* Vol. 107, no. 311, NS2007-103, pp. 53–57, November 2007. (in Japanese)

4. Yuta Yamamoto, Nobuhiko Itoh and Miki Yamamoto, "Performance evaluation of window size in proxy-based TCP for multi-hop wireless networks," *IEICE*

Technical Report, Mobile Multimedia Communication, Vol. 108, no. 44, MoMuC2008-19, pp. 109–114, May 2008. (in Japanese)

5. Yuta Yamamoto, Nobuhiko Itoh and Miki Yamamoto, "Performance improvement of proxy-based TCP in ad hoc networks,"*IEICE Technical Report, Network System,* Vol. 108, no. 203, NS2008-49, pp. 39–44, September 2008. (in Japanese)

6. Nobuhiko Itoh, Yasuhiro Yamasaki and Hideyuki Shimonishi, "An efficient calculation of passive type loop detection method and it's implimentation on OpenFlow-based network," *IEICE Technical Report, Network System,* vol. 109, no. 448, NS2009-167, pp. 31–36, March 2010. (in Japanese)

7. Nobuhiko Itoh, Ippei Akiyoshi and Yasuhiro Mizukoshi, "A study on mobile network architecture for distributed contents," *Proceedings of the 2012 IEICE General Conference,* Vol. 2 p. 54, March 2012. (in Japanese)

8. Nobuhiko Itoh, Ippei Akiyoshi and Yasuhiro Mizukoshi, "A Study on Flow Entry Reduction Scheme for Fast Path Switching," *Proceedings of the 2012 IEICE Society Conference,* Vol. 2 p. 10, September 2012. (in Japanese)

9. Ippei Akiyoshi, Nobuhiko Itoh and Yasuhiro Mizukoshi, "A Study on Network Control for Virtualization of Mobile Core Networks," *Proceedings of the 2012 IEICE Society Conference,* Vol. 2 p. 9, August 2012. (in Japanese)

10. Yasuhiro Mizukoshi, Nobuhiko Itoh and Ippei Akiyoshi, "A Study on Dynamic Traffic Control of Mobile Network after a Devastating Disaster," *Proceedings of the 2012 IEICE Society Conference,* Vol. 2 p. 7, August 2012. (in Japanese)

11. Nobuhiko Itoh, Jun Nishioka, Ippei Akiyoshi and Yasuhiro Mizukoshi, “Flow Entry Reduction Scheme Considering Path Switching for OpenFlow-based Networks,” *IEICE Technical Report, Network System,* Vol. 113, no. 388, NS2013-171, pp. 47–52, January 2014. (in Japanese)

12. Nobuhiko Itoh and Eiji Kawai, “A Study on Flow Entry Reduction Scheme for OpenFlow Networks,” *Proceedings of the 2014 IEICE Society Conference,* Vol. 2 p. 29, September 2014. (in Japanese)

13. Nobuhiko Itoh, Shuji Ishii, Shuichi Saito, Yoshihiko Kanaumi, Eiji Kawai and Shinji Shimojo, “RISE Orchestrator for Wide-Area SDN Testbed,” *Proceedings of IPSJ SIG Internet and Operation Technology Symposium 2014,*, pp. 9–16, November 2014. (in Japanese)

14. Nobuhiko Itoh, Shuichi Saito, Yoshihiko Kanaumi, Eiji Kawai, Shinji Shimojo and Kazuya Suzuki, “Study on RISE architecture for wide-area deployment,”*IEICE Technical Report, Network System,* Vol. 115, no. 251, NS2015-108, pp. 109–114, October 2015. (in Japanese)

15. Takanori Iwai, Akihito Kohiga and Nobuhiko Itoh, “[Invited] Introduction of mobile network research toward IoT era,” *Annual Conference 2016, Academic exchange for information environment and strategy (AXIES)* December 2016. (in Japanese)

16. Takanori Iwai, Akihito Kohiga and Nobuhiko Itoh, “[Invited] Direction of Mobile Network Architecture and Context-Aware Network Control Technology towards IoT era,” *IEICE Technical Report, Communication Quality,* Vol. 116, no. 403, CQ2016-95, pp. 31–34, January 2017. (in Japanese)

17. Takanori Iwai, Akihito Kohiga, Nobuhiko Itoh and Seiichi Koizumi, "[Invited] Context-Aware Network Control Technology in Mobile Network for IoT Services," *IEICE Technical Report, Network System,* Vol. 116, no. 382, NS2016-128, pp. 49–52, December 2017. (in Japanese)

18. Nobuhiko Itoh, Hiroya Kaneko and Takanori Iwai, "Packet Scheduling Method Considering Wireless Channel Quality and Deadline in Mobile Networks," *Proceedings of the 2017 IEICE General Conference,* Vol. 2 p. 315, March 2017. (in Japanese)

19. Akihito Kohiga, Takanori Iwai, Hiroya Kaneko and Nobuhiko Itoh, "Application response time regulation system for auto mobility system," *IPSJ SIG Technical Report, System Software and Operating System,* Vol. 2017-OS-139, no. 4, pp. 1–13, February 2017. (in Japanese)

20. Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Network resource control for guaranteeing IoT application deadline in mobile networks," *IEICE Technical Report, Communication Quality,* Vol. 117, No. 68, CQ2017-16, pp. 13–18, May 2017. (in Japanese)

21. Motoki Morita, Nobuhiko Itoh and Takanori Iwai, "Deadline-Aware Scheduling Scheme for Realizing Connected Car in Mobile Networks," *IEICE Technical Report, Communication Quality,* Vol. 117, No. 486, CQ2017-128, pp. 141–146, March 2018. (in Japanese)

22. Takanori Iwai, Ryogo Kubo, Nobuhiko Itoh and Kozo Satoda, "A bandwidth-reserved transmission rate control method for networked machine control systems," *IEICE Technical Report, Communication Quality,* Vol. 118, No. 71,

CQ2018-19, pp. 7–12, May 2018. (in Japanese)

23. Nobuhiko Itoh, Takanori Iwai, Kozo Satoda and Ryogo Kubo, "A deadline-aware scheduling scheme for vehicle collision avoidance using LTE networks with quality fluctuation," *IEICE Technical Report, Communication Quality,* Vol. 118, No. 192, CQ2018-46, pp. 1–6, August 2018.

24. Takanori Iwai, Nobuhiko Itoh and Kozo Satoda, "Adaptive network control technology towards IoT era," *Proceedings of the 2018 IEICE Society Conference,* BT-2-1, September 2018. (in Japanese)

25. Nobuhiko Itoh, Takanori Iwai and Ryogo Kubo, "Congestion and deadline-aware scheduling of LTE networks for connected car applications," *IEICE Technical Report, Communication Quality,* Vol. 119, No. 183, CQ2019-57, pp. 1–6, August 2019. (in Japanese)

Awards

1. Best Paper Award

   The IEEE International Communications Quality and Reliability Workshop (CQR 2017)

   Date: May 16th, 2017

2. Distinguished Contributions Award

   IEICE Communications Society

   September 11th 2019.

3. Best Research Award

   Technical Committee on Communication Quality, IEICE Communications So-

ciety

Date: July 16th, 2020.

Committees

1. 2017-2019

   Assistant Secretary, Technical Committee on Network Virtualization, the IEICE Communications Society

2. 2017

   Member, Organizing and Program Committee, the 7th International Symposium on Network Virtualization, the IEICE Communications Society

3. 2018

   Member, Organizing and Program Committee, the 8th International Symposium on Network Virtualization, the IEICE Communications Society

4. 2019

   Special Session Organizer, the IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW 2019)

5. 2019

   Member, Technical Program Committee, the 2nd World Symposium on Communication Engineering (WSCE 2019)

6. 2020

   Member, Technical Program Committee, Communication QoS, Reliability and Modeling Symposium, the IEEE International Conference on Communications 2020 (ICC 2020)