NTT DOCOMO

Technical Journal

Technical Journal
Vol.23 No.3 | Jan. 2022

NTT docomo

Changes in Standardization Activities Brought About by COVID-19

DOCOMO Communications Laboratories Europe GmbH
Goro Kunito

It is generally felt that the COVID-19 pandemic has reaffirmed the importance of the communications network as an infrastructure. In the communications network, activities referred to as “standardization” [1] are essential for creating common protocols for transferring information between terminals of different manufacturing companies, between network devices, etc. DOCOMO Communications Laboratories Europe GmbH (DOCOMO Euro-Labs) is engaged in activities involving the standardization of 5G evolution and future 6G at various standardization-related organizations such as the 3rd Generation Partnership Project (3GPP), European Telecommunications Standards Institute (ETSI)*1, 5G Alliance for Connected Industries and Automation (5G-ACIA)*2, and One6G*3. Against this background, I would like to tell you about the effects that this pandemic is having on standardization activities using participation in the 3GPP Service and System Aspects 2 (SA2) meetings held six times a year as an example.

At a 3GPP meeting, proposals (contributions) must be submitted, discussions must be held during the course of the meeting, and a consensus must be reached. It is also important to obtain the approval of other participants and reach a compromise when opinions differ.

Before the pandemic outbreak, different regions around the world such as Europe, North America, and Asia would take turns in holding meetings where delegates from 3GPP member companies from around the world would come together. Such a meeting would last five days from Monday morning to Friday evening and discussions would be held on different topics in three parallel sessions usually running from 8:00 to 20:00 on any given day. For a delegate, the time before the meeting and along the way to the meeting venue would be set aside for preparing a description of one’s contribution and studying countermeasures to others’ contributions. Then, over the course of the meeting, coffee breaks, lunchtime, and dinnertime would be used to solve problems that had been discussed by exchanging opinions and negotiating with delegates of other companies and organizations, and it was common to review and revise the contributions until late at night for the next day. Furthermore, in parallel with these activities, a contribution might be studied together with a non-traveling member who can offer support from the office at home. It was not unusual to hold discussions with others or to revise contributions right up to the final day and final hour and then go straight to the airport.

After the pandemic outbreak, it was decided to replace face-to-face meetings with an “e-meeting” format in which discussions and debates would be held by e-mail over a period of 10 (or 5) business days.

Date	10	17	25	26	27	28
Day	Mon	Mon	Tue	Wed	Thu	Fri
Time (UTC)	23:59	00:00	16:00	16:00		17:00
	Contribution submission deadline	e-meeting begins	Revision deadline	Final comments deadline		e-meeting ends

Figure 1 Schedule example

Figure 1 shows an example of a schedule for such an e-meeting held over 12 days (10 business days).

After the start of the e-meeting and up to the revision deadline, each delegate will make comments on submitted contributions while contributors will respond to comments and update and resubmit their contributions. This will be followed by discussions on whether to approve or disapprove a contribution as needed up to the final comments deadline. In addition, non-business days such as weekends will be a time for each delegate to rest up without exchanging e-mails or updating contributions.

The following problems have arisen on changing the meeting format to an e-meeting.

- When holding traditional face-to-face meetings, time constraints limited the number of contributions that could be handled and the chairperson would request that submitted contributions be carefully screened and selected. In e-meetings, however, the number of contributions has increased by about 25% on average.
- In e-meetings, e-mails fly back and forth 24 hours a day with discussions held in parallel with the result that the number of e-mails over 5 business days during the meeting period has come to about 5,000, or about 10 times the usual number up to now. What used to take several minutes of conversation to settle during a coffee break or other break times in traditional face-to-face meetings is now done by e-mail, which requires time and makes it difficult to hold good discussions.
- Although e-meetings make it unnecessary to travel to a venue, the above problems have been accompanied by a need for human and temporal resources.

Despite these conditions, those individuals involved in global standardization activities are working together on a daily basis with the aim of achieving an even better communications network of the future.

NTT DOCOMO has been compiling a variety of use cases and technical concepts expected of 5G evolution and 6G in a white paper since before the outbreak of the COVID-19 pandemic [2]. The changes in our view of society and the world brought about by this pandemic have the potential of redefining our social, industrial, and economic mechanisms and giving birth to new business opportunities and use cases [3]. As a member of NTT DOCOMO R&D, we would like to continue making contributions to standardization activities with the aim of achieving new and innovative business opportunities and use cases.

REFERENCES

[1] M. Sumita et al.: “Basic Knowledge and Practical Techniques of International Standardization Activities, Part 2 Standardization in Practice Edition, Flow of 1st Standardization Meeting and Participation Guidelines,” NTT DOCOMO Technical Journal, Vol.14, No.2, pp.61–68, Jul. 2006 (In Japanese).

[2] NTT DOCOMO: “White Paper: 5G Evolution and 6G.” https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_v3.0.pdf

[3] H. Morikawa: “How to Face Beyond 5G and 6G,” The Journal of the Institute of Electronics, Information and Communication Engineers, Vol.104, No.5, pp.394–397, May 2021 (In Japanese).

*1 ETSI: The standardization organization concerned with telecommunications technology in Europe.

*2 5G-ACIA: A global alliance studying the application of 5G technology to industry-oriented use cases such as factory automation.

*3 One6G: An association formed on the basis of academic-industrial cooperation to promote the creation and development of 6G-related technology.

NTT DOCOMO
Technical Journal
Vol.23 No.3 Jan. 2022

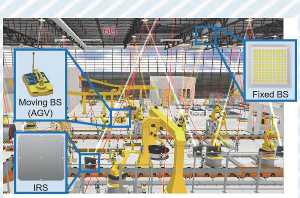
[Contents]

DOCOMO Today

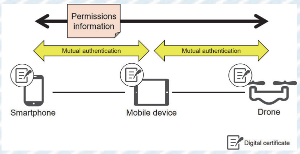
Technology Reports (Special Articles)

Special Articles on 5G evolution & 6G (2)

(P.5)



(P.13)

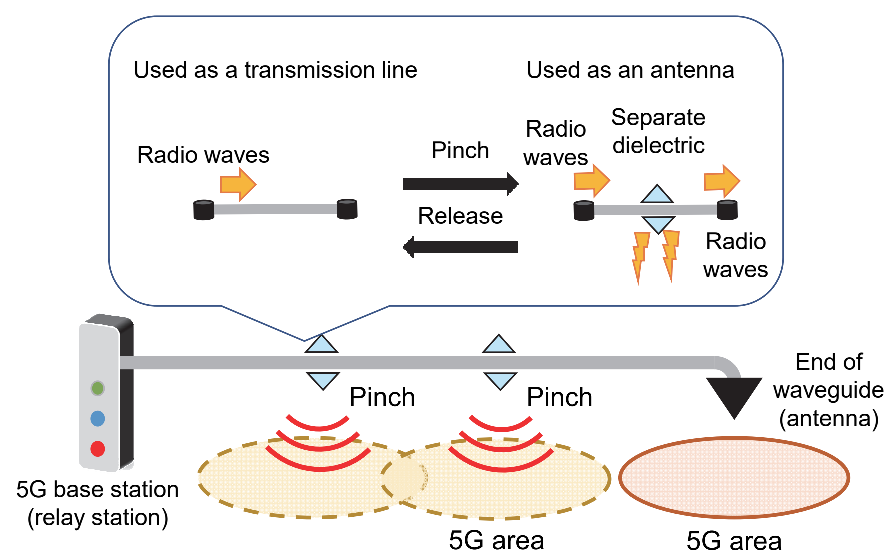


(P.27)

Technology Reports

(P.52)

News



Technology Reports (Special Articles) Pinching Antenna—Using a Dielectric Waveguide as an Antenna— (P.5)
Deploying a communication area with a bended antenna

Technology Reports (Special Articles)
Millimeter-wave Communication Area Deployment | Dielectric Waveguide | Antenna
Special Articles on 5G evolution & 6G (2)

Pinching Antenna —Using a Dielectric Waveguide as an Antenna—

6G-IOWN Promotion Department Atsushi Fukuda Hiroto Yamamoto
Hiroshi Okazaki Yasunori Suzuki
Radio Access Network Development Department Kunihiro Kawai

We propose a method for easily and economically deploying and expanding communication areas in high-frequency bands (millimeter-wave bands) that are now being used in the 5G era and studied for use in the 6G era. This method radiates radio waves from part of a dielectric waveguide—a transmission medium for high-frequency bands—to make the surrounding area a communication area. The ability to radiate radio waves from any point along a dielectric waveguide makes it possible to quickly deploy a variety of communication areas depending on the place and environment.

1. Introduction

In Japan, commercial services for the 5th Generation mobile communications system (5G) were launched in March 2020. As a mobile communications system, 5G marked the first practical use of wideband, high-speed communications using a (quasi-) millimeter-wave band, in particular, the 28-GHz frequency band [1]. The 28 GHz band and higher bands are expected to see expanded use in the future.

In high-frequency bands, it is known that propagation loss*1 is generally high [2]. Furthermore, in a Non-Line-Of-Sight (NLOS)*2 communication environment between the base station and terminal, high-frequency radio waves are more susceptible to electromagnetic shielding than low-frequency radio waves. This creates concern that receiving power at the terminal will dramatically drop and that the communication area will be narrower than that with low frequency bands. Consequently, for 5G and the 6th Generation mobile communications system (6G) for which the use of high-frequency bands of 28 GHz and higher are being studied, the need is felt for securing a Line-Of-Sight (LOS)*3 communication environment between the base station and terminal when pursuing high-speed, large-capacity wireless communications using high-frequency bands [3].

For example, the demand for IoT within factories is expected to grow from here on, and there are great expectations for using 5G to meet this demand from the viewpoints of low-latency and large-capacity transmission. On the other hand, since there are many cases of electromagnetic shielding caused by people and machines in a factory, an NLOS environment tends to form, and on top of this problem, such electromagnetic shielding may move and even factory layout modifications may occur due to changes in the production line. To deal with such fluctuations in the communication environment, a method is needed to quickly and economically provide LOS communication environments.

One method for ensuring a LOS communication environment is to distribute the arrangement of base-station antennas while controlling the directivity*4 of the radio waves radiated from a base-station antenna. This method, in addition to ensuring a LOS communication environment, can be expected to improve spectrum efficiency*5 and communication capacity over the entire system [3]. However, the method is difficult to apply if antenna installation sites and power-supply paths are difficult to set up, and an increase in installation costs accompanying an increase in the number of distributed antennas can also be a problem.

The method proposed in this article for ensuring a LOS communication environment leaks radio waves from part of a dielectric waveguide used as a transmission medium for high-frequency bands and uses these radiated radio waves to deploy a communication area surrounding that point. A waveguide that leaks radio waves is called a "leaky waveguide," which is similar to a Leaky Coaxial Cable (LCX)*6 used for deploying communication areas in underground shopping complexes, tunnels, etc. As such, a leaky waveguide is said to be a high-frequency version of LCX, which suffers from high loss in high-frequency bands. A leaky waveguide also provides new uses not possible with LCX. In the case of LCX, radio waves are leaked from gaps set beforehand in the outer conductor of the coaxial cable, which means that once a LCX is laid, it continues to leak radio waves from those places. On the other hand, the proposed leaky waveguide enables the places for leaking radio waves to be controlled as desired. In this article, we outline "bended antenna" and "pinching antenna" that embody the proposed leaky waveguides and describe verification experiments using a 60 GHz video transmission system.

2. Using a Waveguide as an Antenna

2.1 Waveguides

Media for transmitting signals in high-frequency

*1 Propagation loss: The amount of attenuation in the power of a signal emitted from a transmitting station until it arrives at a reception point.

*2 NLOS: Describes an environment where there are obstacles between the transmitter and receiver. In this case, communication can only take place over waves that have been reflected, refracted, etc.

*3 LOS: Describes an environment where there are no obstacles between the transmitter and receiver, allowing them to communicate via direct waves.

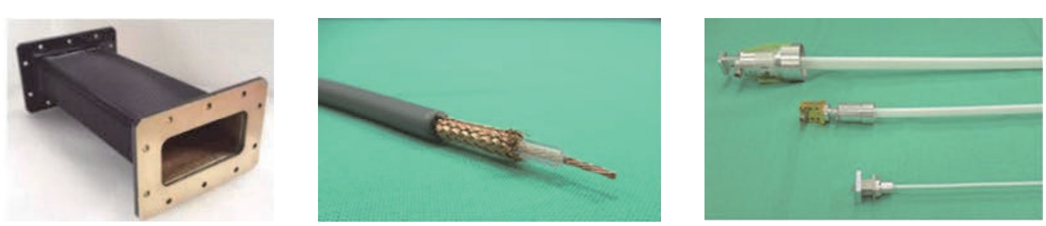
*4 Directivity: An index of direction and intensity of a radio wave radiated in space.

bands include (a) metallic waveguides, (b) coaxial cables, and (c) dielectric waveguides (**Photo 1**).

(a) A metallic waveguide has a cross-sectional shape in which a conductor surrounds a dielectric[*7] or air. It is used in satellite earth stations and other applications. Here, the cross-sectional shape of a metallic waveguide determines the target frequency band, and high-frequency radio waves can be transmitted in this way. However, the cross section for transmitting low-frequency waves is large. For this reason and the fact that a metallic waveguide cannot be easily bent, constraints may arise in the handling of this type of waveguide.

(b) A coaxial cable has a cross-sectional shape in which the center conductor and outer conductor are arranged in a concentric manner. It is used widely from direct-current to high-frequency-band applications. It is common to fill the space between both conductors with a dielectric. There is little radio-wave leakage from the outer conductor, and a thin cable features a certain amount of flexibility, but loss in high-frequency bands is higher than that in metallic waveguides, so coaxial cables are mainly used in microwave-and-lower frequency bands. A coaxial cable with gaps set in the outer conductor is called an LCX, and having radio waves radiate from those gaps makes it possible to deploy communication areas in zones normally out of reach of radio waves such as underground shopping complexes and tunnels.

(c) A dielectric waveguide has a structure in which a bar-shaped dielectric is surrounded by a dielectric of different permittivity[*8]. Making the permittivity of the inner dielectric higher than that of the waveguide periphery makes it possible to confine and transmit high-frequency radio waves mainly through the inner dielectric. The dielectric waveguide targeted in this article uses PolyTetraFluoroEthylene (PTFE)[*9] with a relative permittivity[*10] of 2.1 for the inner conductor and air with a relative permittivity of nearly 1.00 for the outer conductor. Here, radio waves suffer loss when propagating through the dielectric, and while this loss generally increases as the frequency becomes higher, this increase is smaller than that of a coaxial cable, so dielectric waveguides are used as a transmission medium in high-frequency bands.



(a) Metallic waveguide (b) Coaxial cable (c) Dielectric waveguide

Photo 1 Examples of media for transmitting signals in high-frequency bands

---

*5 **Spectrum efficiency:** The number of data bits that can be transmitted per unit time over a particular frequency spectrum.

*6 **LCX:** A coaxial cable having a mechanism for leaking radio waves from gaps set in its outer conductor.

*7 **Dielectric:** An insulator in which direct current cannot flow. Insulators include plastics and other materials.

*8 **Permittivity:** An index of the ease of polarizing a dielectric.

*9 **PTFE:** A fluororesin consisting of fluorine and carbon atoms.

*10 **Relative permittivity:** An index expressing the ratio of a material's permittivity when taking permittivity in a vacuum to be 1.

To establish a means of ensuring a LOS communication environment for high-frequency bands, NTT DOCOMO has been studying a configuration that minimizes the propagation distance and prevents the generation of NLOS communication environments as much as possible by using the low-loss transmission characteristics of dielectric waveguides in high-frequency bands and radiating radio waves in the vicinity of terminals. Up to now, two types of antennas have been proposed in relation to radiating waves from a dielectric waveguide, that is, to using a dielectric waveguide as an antenna: the bended antenna that bends the dielectric waveguide and radiates radio waves from the bent section, and the pinching antenna that pinches the dielectric waveguide with a separate dielectric and radiates radio waves from that location. The following describes these two types of antennas.

2.2 Bended Antenna

The phenomenon of radiating radio waves to the outside occurs at the bent section of a dielectric waveguide. This radiation is accompanied by a corresponding amount of power loss in the waveguide. However, the primary objective of a dielectric waveguide is to deliver radio waves to a distant location with as little loss as possible. Dielectric waveguides have therefore been used in a manner that avoids bending. On the other hand, a bended antenna proactively uses this radiation phenomenon at the bent section of a dielectric waveguide as shown in **Figure 1** [4]. Since the dielectric waveguide can be structured with easy-to-bend materials, it can be bent at a location where radiation is desired as shown in the figure. In other words, radio waves can be radiated at any location along the waveguide to deploy a communication area at that point. In addition, when removing such bending and straightening out the waveguide, this radiation phenomenon ceases to exist, which means that the radio-wave radiation can be terminated at any time and the corresponding communication area removed. On the other hand, since a certain length of waveguide is needed to

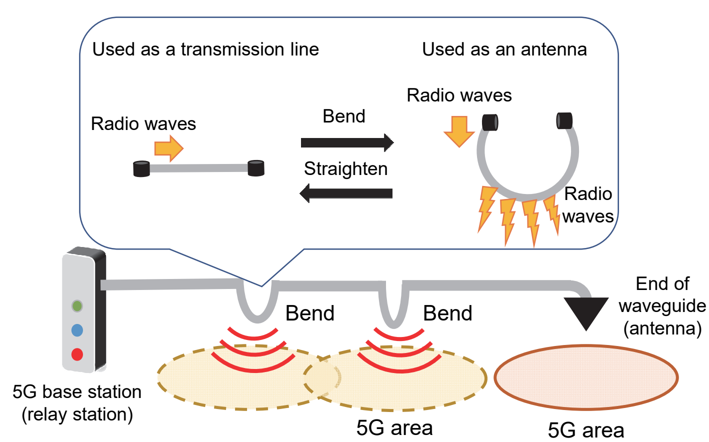


Figure 1 Deploying a communication area with a bended antenna

configure the bent sections for forming bended antennas, a spare amount of length is needed when laying out the waveguide. Attention must also be given to deterioration in waveguide characteristics due to repeated bending and extension.

2.3 Pinching Antenna

Adding a separate dielectric to the dielectric waveguide enables a portion of the radio waves propagating along the waveguide to be induced into that dielectric. As shown in **Figure 2**, a pinching antenna applies the phenomenon of radio-wave radiation from the separate dielectric to form the antenna. Similar to a bended antenna, a pinching antenna allows pinching locations to be selected as desired, which means that radio waves can be radiated at any location along the waveguide to form a communication area as shown in the figure. Here as well, this radiation phenomenon ceases to exist when releasing the pinch so that radio-wave radiation can be terminated at any time and the corresponding communication area removed. It is also possible for a pinching antenna to receive signals from the outside at its radiation (pinching) locations. A pinching antenna is also easy to install since it only involves the attaching of separate dielectrics at some points along the dielectric waveguide.

3. Verification Experiment of Dielectric Waveguide Antenna

We performed a verification experiment to test the radio-wave radiation phenomenon of a pinching antenna. The experiment used High-Definition Multimedia Interface (HDMI) transmitter/receiver units commercially available from Sharp Corporation (VR-WH1). Main radio specifications are listed in **Table 1.** This product consists of wireless transmitter/receiver units conforming to the Wireless HD standard using the 60 GHz band. The transmitter converts the input HDMI signal into

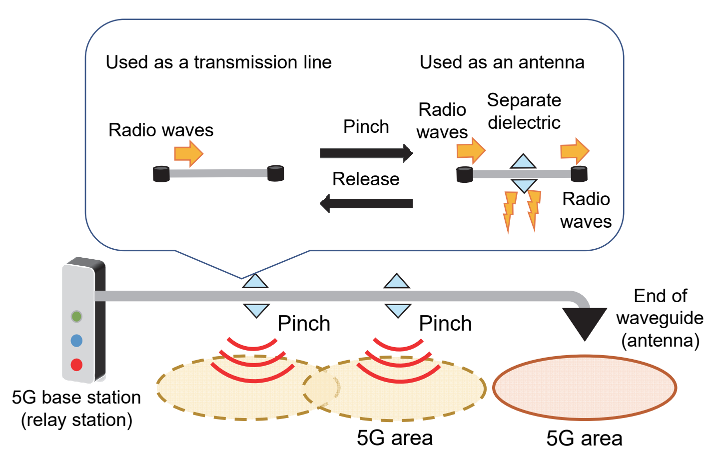


Figure 2 Deploying a communication area with a pinching antenna

radio waves and transmits them using a wide bandwidth (1.76 GHz) in the 60 GHz band, and the receiver generates HDMI signals from received radio waves. In the experiment, spatially radiated radio waves are input into the dielectric waveguide from the transmitter and then radiated from the pinching antenna and received at the receiver.

3.1 Video Transmission Experiment Using a Pinching Antenna

A block diagram of the experimental system is shown in **Figure 3**. In this system, a Personal Computer (PC) plays back the video and converts it into an HDMI signal. A splitter then divides the output HDMI signal inputting one signal into the transmitting monitor and the other signal into the transmitter. The transmitting monitor is used to verify the video played back on the PC. The transmitter, meanwhile, converts the HDMI signal into 60 GHz radio waves and inputs them into the dielectric waveguide. The receiver then inputs the radio waves radiated from the pinching antenna and converts them into an HDMI signal. Finally, the receiver outputs the HDMI signal into the receiving monitor so that the received video can be verified. In addition, the separate dielectric of the pinching antenna is attached to the tips of an ordinary off-the-shelf clothespin.

The states of the transmitting monitor and receiving monitor before and after pinching with the

Table 1 Main radio specifications of wireless HDMI transmitter/receiver units

Compliant standard	Wireless HD 1.1
Transmission system	HRP/LRP
Center frequency	60.48 GHz (Ch2) 62.64 GHz (Ch3)
Bandwidth	1.76 GHz

HRP: High Rate PHY
LRP: Low Rate PHY

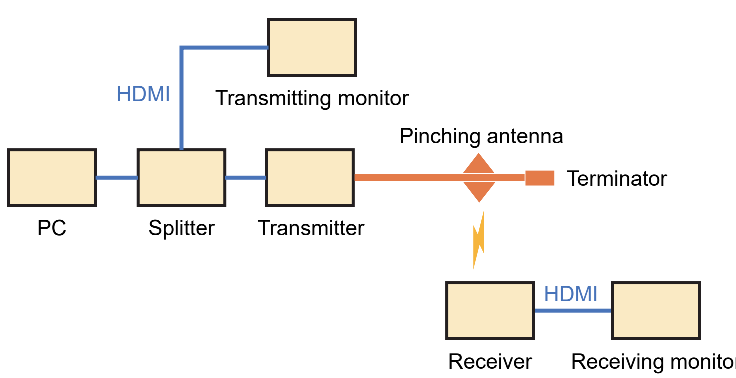
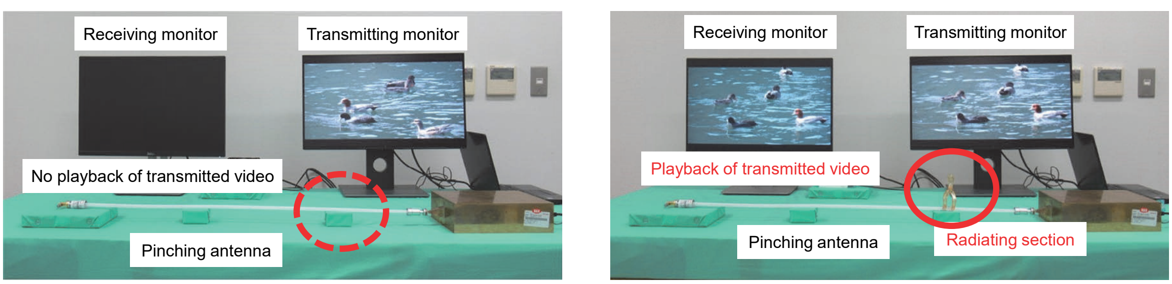


Figure 3 Block diagram of experimental system

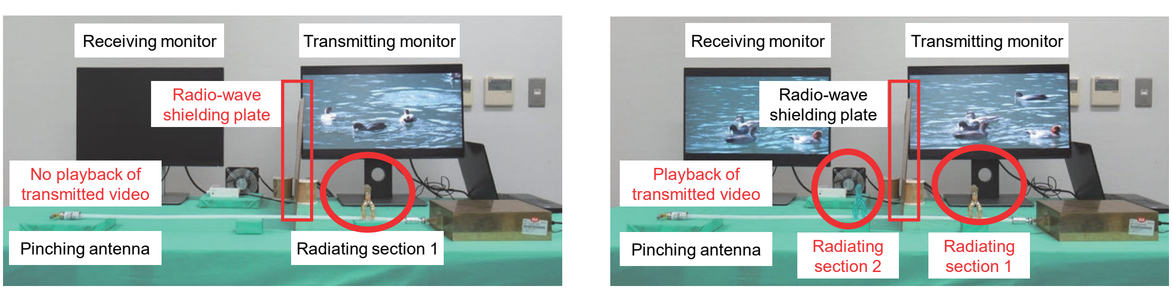
clothespin are shown in **Photo 2** (a) and Photo 2 (b), respectively. From Photo 2 (b), it can be seen that pinching the dielectric waveguide with the clothespin attaches the separate dielectric to the waveguide causing radio waves to be radiated from the separate dielectric, and that receiving those radiated waves at the receiver results in playback of the same video as the transmitting monitor on the receiving monitor connected to the receiver. From Photo 2 (a), it can be seen that no pinching with the clothespin results in no radiation and no reception of radio waves so that nothing is displayed on the receiving monitor.

Photo 3 shows the transmitting monitor and receiving monitor when placing a radio-wave shielding plate between the radiating section and receiving section after pinching with the clothespin as shown in Photo 2 (b). The shielding plate loses a LOS environment between the radiating section and receiving section resulting in no reception of radio waves and no display on the receiving monitor as shown in Photo 3 (a). In such a situation, pinching the dielectric waveguide at any other location enables radio waves to be radiated (the dielectric waveguide may be pinched at multiple locations). As shown in Photo 3 (b), pinching the dielectric waveguide at the location where a LOS environment can be created between another radiating section and the receiving section enables radio waves to be received and the same video to be displayed on the receiving monitor and transmitting monitor. If the shielding plate happens to move, changing the pinching location makes it easy to create another new LOS environment.



(a) Separate dielectric removed (before pinching)　　(b) Separate dielectric attached (after pinching)

Photo 2 Video transmission experiment through radio-wave radiation by pinching antenna



(a) Shielding plate installed (one radiating section)　　(b) Shielding plate installed (two radiating sections)

Photo 3 Avoiding radio-wave shielding with a pinching antenna

4. Conclusion

This article explained how a dielectric waveguide could be used as a transmission medium for high-frequency bands, described its application to antennas, and presented the bended antenna and pinching antenna as means of flexibly deploying communication areas. Going forward, we plan to perform trials of deploying communication areas in real-world environments.

REFERENCES

[1] Y. Kojo et al.: "Overview of 5G Commercial Service," NTT DOCOMO Technical Journal, Vol.22, No.1, pp.4–9, Jul. 2020.

[2] T. Imai et al.: "Radio Propagation for 5G," NTT DOCOMO Technical Journal, Vol.17, No.4, pp.40–49, Apr. 2016.

[3] NTT DOCOMO: "White Paper: 5G Evolution and 6G (Version 3.0)," Feb. 2021. https://www.nttdocomo.co.jp/english/corporate/technology/whitepaper_6g/

[4] K. Kawai, T. Takada, A. Fukuda, H. Okazaki and S. Narahashi: "A new area formation approach for millimeter wave communication systems employing a dielectric waveguide," Proc. of the 45th European Microwave Conference, pp.1088–1090, Sep. 2015.

Technology Reports (Special Articles) 6G 100 GHz Band System-level Simulator

Special Articles on 5G evolution & 6G (2)

6G System-level Simulator —Toward 100 GHz Band, 100 Gbps Extreme-high-data-rate Communications—

6G-IOWN Promotion Department Tatsuki Okuyama Satoshi Suyama
Nobuhide Nonaka Takahiro Asai

With 6G, the target is to add the 100 – 300 GHz bands to the millimeter-wave band introduced by 5G, which means that we can expect extreme-high-data-rate communications beyond 100 Gbps as a peak data rate using so-called terahertz waves. At NTT DOCOMO, we have already begun research and development toward the use of terahertz waves in 6G with the aim of conducting trials using experimental equipment in the future. However, at this initial stage of study, the possibility of improving system performance using terahertz waves to achieve extreme-high-data-rate communications of 100 Gbps must first be demonstrated through performance evaluations based on simulations. With this in mind, we have developed a real-time system-level simulator using the 100 GHz band and used this simulator to show that a throughput in excess of 100 GHz per user can be achieved in an environment with multiple users for two types of indoor scenarios. This article describes these simulations in detail.

1. Introduction

The commercial deployment of the 5th Generation mobile communications system (5G) has been progressing throughout the world, but the penetration of 5G services in society has uncovered issues with

5G and generated expectations of further evolution. In 5G, services have been rolled out using the Sub6*1 band and the 28 GHz millimeter-wave*2 band. With regard to the latter, a 400 MHz wide bandwidth has been allocated in Japan, which can provide a high data rate compared with the 100 MHz bandwidth of the Sub6 band.

At the same time, opportunities for using high-frequency bands toward the 6th Generation mobile communications system (6G) have recently been increasing. For example, the Federal Communications Commission (FCC) in the United States is opening up the 95 GHz – 3 THz frequency bands higher than the millimeter-wave bands for 6G trials. The study of wireless technologies and the development of devices toward the use of such high-frequency bands are moving forward [1] – [3]. However, to achieve this objective, the possibility of improving system performance through the use of high-frequency bands must be clarified early and performance evaluations in real environments using experimental equipment must be performed as research and development advances. However, equipment development takes time, so it is not that easy to perform evaluations in real environments at the initial study stage. It is therefore necessary to demonstrate performance by simulations to determine, for example, the level of communication speeds that can be achieved by using terahertz waves*3. To this end, NTT DOCOMO has developed a system-level simulator*4 for the purpose of clarifying the feasibility of extreme-high-data-rate communications by 6G in near-real-world environments. With this simulator, we targeted two types of scenarios, namely, a shopping mall and factory, each featuring both stationary and moving Mobile Stations (MSs). We evaluated, in particular, the downlink throughput per MS using the 100 GHz band taken from the 95 GHz – 3 THz frequency bands mentioned above.

In this article, we show by simulations that a throughput of 100 Gbps per MS can be achieved by ultra-broadband*5 transmission (bandwidth = 8 GHz), which marks a step toward the practical use of 6G. We also describe how using multiple units of a drone Base Station (BS) and Intelligent Reflecting Surface (IRS)*6 enables the number of MSs capable of 100 Gbps throughput to be increased while compensating for high propagation loss*7 in the 100 GHz band.

2. Adoption of High-frequency Bands toward 6G and New Wireless Communication Technologies

In 6G, we can envision new combinations of requirements beyond the high-data-rate/high-capacity, low-latency, and massive-connectivity features of 5G as well as use cases that require extreme performance difficult to achieve in 5G [3]. Specifically, this will mean improving communication speeds even further in 6G through “extreme-high-data-rate and high-capacity communications” beyond 100 Gbps to achieve new sensory services that are comparable or even superior to the quality of experiences achieved through the five human senses. We can also expect “extreme coverage extension” that will cover areas such as the sky, sea, and space not

---

*1 Sub6: A division of the frequency band. A radio signal with a frequency between 3.6 GHz and 6 GHz.

*2 Millimeter waves: A division of the frequency band. A radio signal with a frequency between 30 GHz and 300 GHz.

*3 Terahertz waves: Electromagnetic waves with a frequency of around 1 THz. Often used to refer to frequencies ranging from 100 GHz to 10 THz.

*4 System-level simulator: In contrast to link-level simulation that simulates the behavior between base stations and terminals, system-level simulation is an evaluation method that, within an environment containing multiple base stations and terminals, incorporates quality control based on the selection of communication terminals and on the radio-wave propagation environment.

*5 Ultra broadband: Bandwidth of 100 MHz or greater. In Japan, 400 MHz of bandwidth has been assigned in the 28 GHz band for 5G radio communications.

*6 IRS: A reflecting surface that can arbitrarily design the direction of reflected waves and beam shape by a two-dimensional arrangement of structures called metamaterial that are microscopically small with respect to wavelength.

presently covered by current mobile communications systems, "extreme low power consumption and cost reduction" per bit in the future adoption of millimeter waves and terahertz waves, "extreme low latency" achieving End-to-End (E2E) communications under 1 ms, and "extreme reliable communication" for achieving high reliability and high security. We can also envision "extreme massive connectivity" of 10 million devices per square kilometer and the provision of functions for sensing the real world as in ultra-high-accuracy positioning of MSs on the order of centimeters using radio waves of mobile communications [4].

In 6G, moreover, the use of terahertz waves will enable signal bandwidths dramatically wider than those of 5G, so there are expectations of extreme-high-data-rate communications in excess of 100 Gbps. On the other hand, terahertz waves have strong straight-line propagation properties and large propagation loss compared with conventional millimeter waves, which prevents long-distance transmission. This is a technical problem that needs to be addressed. It is therefore important to clarify and model signal propagation characteristics*8 in terahertz waves, establish 6G radio access technologies based on those characteristics, and develop wireless device technologies for these high-frequency bands.

It is common in current mobile communications networks to use a fixed network topology that provides area coverage with BSs installed by a telecommunications carrier. In 6G, however, securing coverage in high-frequency bands and improving connectivity will require an even more advanced network configuration called New Radio Network Topology that will include optimal path selection through the cooperation of multiple access points in the vicinity of a MS as well as diversity*9 transmission and reception [3]. However, to achieve New Radio Network Topology, a variety of studies still need to be made. These include the use of existing objects such as streetlamps and traffic lights for mounting communication antennas, the development of advanced radio relay technologies such as Integrated Access and Backhaul (IAB)*10 and repeater*11 equipment for high-frequency bands, IRSs that can dynamically control reflection intensity and directivity*12, inter-MS linking, and moving BSs. In 6G, the development of advanced Massive Multiple Input Multiple Output (Massive MIMO)*13 is an important technology area. Here, the use of even more antenna elements and an increase in the number of streams (many-layer scheme) are progressing, and the use of Massive MIMO in a distributed antenna arrangement called Distributed MIMO*14 in combination with New Radio Network Topology shows promise.

3. 6G System-level Simulator toward Use of High-frequency Bands

3.1 Simulator Overview

To demonstrate the 6G requirements and technical concepts described in an NTT DOCOMO white paper and special article in this journal [3] [4] and to present the possibilities of terahertz waves, it is important, in the end, to perform evaluations in real environments using real equipment.

*7 Propagation loss: The amount of attenuation in the power of the signal emitted from the transmitting station till it arrives at the reception point.

*8 Signal propagation characteristics: Refers to characteristics such as propagation losses, power and delay profiles, and angular profiles.

*9 Diversity: A general name for technologies designed to improve the quality and reliability of communications using MIMO antennas. In particular, non-closed loop types.

*10 IAB: Technology that aims to achieve high-data-rate/high-capacity services over a wide area by applying 5G wireless communications to even backhaul communications and by achieving flexible and low-cost network design and rollout.

*11 Repeater: Relay equipment on the physical layer for amplifying downlink signals received from a base station and forwarding them to a terminal.

*12 Directivity: An antenna radiation characteristic indicating the directional characteristics of radiation strength (or reception sensitivity) from the antenna.

However, the maturing of wireless device technology and equipment development takes time. At the same time, the possibilities that terahertz waves can bring should be tested early, so it is important to evaluate the possibility of achieving extreme-high-data-rate communications in excess of 100 Gbps by computer simulations. In this regard, it has already been shown by link-level simulation*15 that throughputs of 100 Gbps and greater in the 100 GHz band can be achieved [5]. However, to verify the utility of terahertz waves in near-real-world environments, we developed a 6G system-level simulator in the 100 GHz band and used it to evaluate transmission performance.

This simulator applies terahertz waves under the constraints that antenna size is equivalent to that of the 28 GHz band and transmission power is the same as that of 5G. The use of terahertz waves makes it possible to significantly increase the number of antenna elements (hereinafter referred to as "number of elements"), which means that high BeamForming (BF)*16 gain can be obtained. This, in turn, should compensate for the high propagation loss of terahertz waves, which is what the simulator will test. It also implements IRSs and moving BSs to test New Radio Network Topology mentioned above so that any improvement effects with respect to loss caused by shadowing and blocking can be verified.

In this article, we evaluate user throughput and the feasibility of achieving throughput in excess of 100 Gbps in a shopping mall scenario and factory scenario as two examples of indoor environments.

3.2 Simulator Functions

This simulator compares performance between 5G using a millimeter-wave band and 6G using the 100 GHz band. For the two indoor scenarios described above, performance was evaluated by installing multiple fixed BSs at locations determined beforehand. Furthermore, for the 6G performance evaluation, communications by these fixed BSs were supplemented by drone BSs and small-vehicle-type BSs moving along specific paths. Moreover, in addition to communication by personal devices that have been the mainstay up to now, both people and robots were arranged as MSs envisioning the provisioning of services by advanced robots in the future. It was also assumed that a mixture of stationary MSs and MSs moving at walking speed would be present.

Terahertz waves are greatly affected by propagation loss and blocking that are significantly higher than that of millimeter waves. Taking this into consideration, this simulator first calculates propagation loss and attenuation by blocking according to the positional relationship between BSs and MSs and whether blocking exists between those BSs and MSs. It then determines the received power of each MS and uses that received power to determine the MSs communicating with each BS (fixed, drone, moving, IRS). When using an IRS, the simulator calculates attenuation by blocking using the path lengths from BS to IRS and from IRS to MS. Here, to simplify evaluations, the simulator assumes reflection direction by an IRS to be ideally controlled and performs BF by treating the IRS as a virtual fixed BS. Additionally,

---

*13 **Massive MIMO**: MIMO systems transmit radio signals overlapping in space by using multiple antenna elements for transmission and reception. Massive MIMO systems aim to achieve high-speed data communications with greater numbers of simultaneous streaming transmissions while securing service areas. They achieve that aim by using antenna elements consisting of super multi-element arrays to create sharply formed radio beams to compensate for the radio path losses that accompany high-frequency band usage.

*14 **Distributed MIMO**: A MIMO transmission technology that transmits different MIMO streams from multiple base stations to a single mobile station.

*15 **Link-level simulation**: Modeling of the transmitter, receiver, and the physical behavior of the radio propagation path between them, applied in experiments on functionality and performance from transmitter to receiver.

*16 **BF**: A technique for increasing or decreasing the gain of antennas in a specific direction by controlling the amplitude and phase of multiple antennas to form a directional pattern of the antennas.

based on received power calculated from the above, the simulator will switch BSs communicating with certain MSs with no delay.

To provide communications to multiple MSs, a BS schedules the allocation of radio resources*17, generates a transmission weight*18, performs rank control*19, and determines the modulation level*20. It performs scheduling by allocating time in slot units and frequency in Resource Block (RB)*21 units based on a Proportional Fairness (PF)*22 algorithm. It also calculates a PF metric*23 for each MS and RB and allocates resources to the MS-and-RB combination having the largest PF metric. This simulator, however, performs no retransmissions, so retransmission control is not used in this metric calculation.

Then, after resource allocation, the BS determines the modulation level according to the average Signal to Interference plus Noise power Ratio (SINR)*24 estimated within that RB. Furthermore, in each time slot, each BS communicates with one MS and transmits a maximum of eight layers according to the number of layers that can be transmitted. Here, to speed up calculations in system-level simulations, we use a transmission weight that averages by rank number the elements in the column direction of the Hermitian transpose*25 of the propagation channel*26.

The MS, meanwhile, estimates the transmission signal based on the Minimum Mean Squared Error (MMSE)*27 weight using the propagation channel [6]. It then calculates the block error rate and received SINR of this estimated signal and calculates throughput.

3.3 Evaluation Scenarios

This article describes system-level simulations for a shopping mall and factory. These simulations construct a propagation environment based on multipath Rayleigh fading*28.

1) Overview of Shopping Mall Scenario

The shopping mall scenario is shown in **Figure 1**. This environment consists of a three-story structure in which shops are lined up on the sides of an atrium-style mall. For the sake of simplification, the shop section is treated as a wall surface. People and any robots or self-driving vehicles that serve as MSs are either stationary or moving on the 1st floor. Fixed BSs are installed on the walls of the 2nd floor ceiling. In addition, the evaluation for 6G included the installation of IRSs and drone BSs. The IRSs were installed on pillars to mitigate blocking effects caused by those pillars or signboards scattered throughout the mall and were positioned to reflect radio waves of a fixed BS in a desired direction. The drone BSs were made to shuttle back and forth in the air in the lower section of the two-story atrium to provide communications to MSs on the 1st floor. In this evaluation, it was assumed that the backhaul*29 for the drone BSs was ideally constructed.

2) Overview of Factory Scenario

The factory scenario is shown in **Figure 2**. In contrast to the shopping mall scenario, this scenario assumes a large box-shaped environment containing conveyor belts for work and a crane for moving packages as well as Automatic Guided Vehicles (AGVs). In this environment, people, robots, and AGVs are taken to be MSs. Fixed BSs

---

*17 Radio resources: General term for radiocommunication resources (radio transmission power, allocated frequency, etc.).

*18 Transmission weight: A transmission weighting factor for forming a directional pattern by controlling the amplitude and phase of multiple antennas and for increasing/decreasing antenna gain in a specific direction.

*19 Rank control: A method for adaptively changing the number of spatially multiplexed streams according to propagation channel conditions. Given a propagation environment in which the number of eigenspaces (rank) needed for spatial multiplexing is large, the number of spatially multiplexed streams is made large to obtain high throughput.

*20 Modulation level: The number of signal phase points in data modulation. This number is 4 in Quadrature Phase Shift Keying (QPSK) and 16 in 16 Quadrature Amplitude Modulation (16QAM).

*21 RB: A unit of frequency to be allocated when scheduling radio resources.

*22 PF: A technique for allocating radio resources considering fairness among multiple terminals.

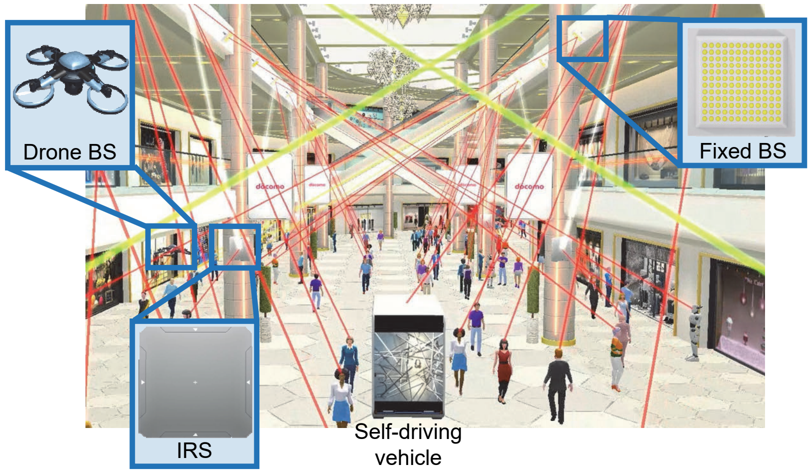


Figure 1 Evaluation environment for shopping mall scenario

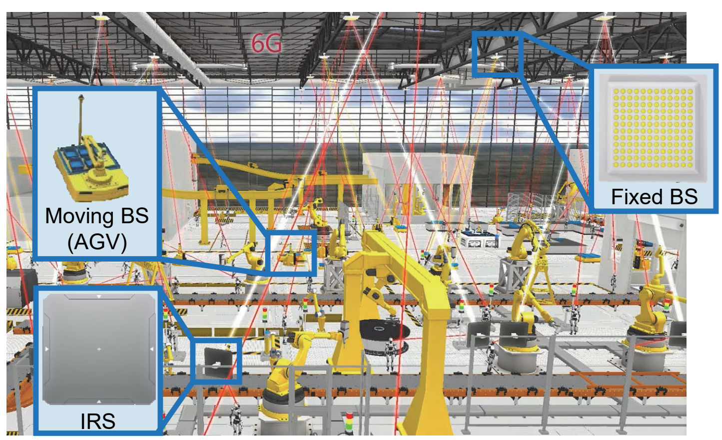


Figure 2 Evaluation environment for factory scenario

are installed on the ceiling and IRSs are installed on walls or pillars so that radio waves can reach MSs positioned at locations that are blocked by the crane or other obstructions. In addition, this environment has no blocking in the ceiling section as in the case of the shopping mall scenario so that radio waves from above are considered to emanate from fixed BSs. Any benefits from drone

*23 Metric: A numerical index. In the PF algorithm, it is the value obtained by dividing the value of the instantaneous communication quality (received power, etc.) by average communication quality over a certain time period.

*24 SINR: The ratio of desired-signal power to the sum of all other interference-signal power and noise power.

*25 Hermitian transpose: The matrix obtained by transposing each element of a matrix or the rows and columns of a complex matrix containing complex numbers and taking the conjugate of each element.

*26 Propagation channel: An individual communication path in wireless communications. In this article, a communication path between transmit/receive antennas.

*27 MMSE: A method for signal computation that minimizes mean square error.

*28 Multipath Rayleigh fading: A phenomenon by which radio signals emitted from a transmit point traverse multiple transmission paths (multi-path transmission) and combine at a moving receive point resulting in severely fluctuating receive levels. It is known that this statistical fluctuation distribution approximates a Rayleigh distribution especially in a non-line of sight propagation environment.

*29 Backhaul: Indicates the route connecting a wireless base station to the core network.

BSs would therefore be difficult to obtain in this scenario. For this reason, some of the AGVs on the floor are treated as moving BSs. The backhaul for these moving BSs operate ideally the same as that for drone BSs.

4. Performance Evaluation of Throughput in Excess of 100 Gbps Using Terahertz Waves

4.1 Simulation Specifications

Simulation specifications are listed in **Table 1**.

Table 1 Simulation specifications

	Shopping Mall		Factory	
Communication system	5G	6G	5G	6G
Center frequency	28 GHz	100 GHz	28 GHz	100 GHz
Bandwidth	400 MHz	8,000 MHz	400 MHz	8,000 MHz
Number of fixed BS elements (vertical × horizontal × sub-arrays)	392 (7×7×8)	4,608 (24×24×8)	128 (4×4×8)	1,152 (12×12×8)
Number of drone BS elements (vertical × horizontal × sub-arrays)	–	1,024 (8×8×8)	–	–
Number of moving BS elements (vertical × horizontal × sub-arrays)	–	–	–	228 (6×6×8)
Number of BSs	10	Fixed: 10, 20 Drone: 4 IRS: 12	12	Fixed: 12, 25 Moving: 3 IRS: 12
BS total transmission power	30 dBm	Fixed: 30 dBm Drone: 15 dBm	30 dBm	Fixed: 30 dBm Moving: 15 dBm
BS element interval	0.5 λ			
BS element gain	5 dBi			
Number of MS elements	32 (omni-directional antenna)			
Number of MSs	100 (moving: 70, stationary: 30)		100 (moving: 50, stationary: 50)	
MS element gain	0 dBi			
Moving speed	3 km/h			
Channel estimation	Ideal			
Maximum number of layers	8			

For 5G, the center frequency*30 and bandwidth were set to 28 GHz and 400 MHz, respectively, while for 6G, the simulator performed ultra-wideband transmission with a bandwidth of 8,000 MHz at a center frequency of 100 GHz. The element interval was set to 0.5 λ for both 5G and 6G. At this time, the antenna panel size for fixed BSs was assumed to be the same for both 5G and 6G with 6G having approximately 10 times more elements than 5G. In addition, a BS is structured with 8 sub-arrays*31 in one flat array*32 with each sub-array forming one analog beam. Beam direction at each BS was oriented to one MS unit and each BS performed Single User MIMO (SU-MIMO)*33 accommodating only one MS. Transmission power was fixed regardless of the number of elements and set to 30 dBm for fixed BSs and 15 dBm for drone and moving BSs. Each MS used an omni-directional antenna*34 having 32 elements with no gain, and MS gain was taken to be the diversity gain achieved by multiple elements. In addition, 100 MSs were used in the evaluation environment, and among this number, 70 in the shopping mall and 50 in the factory were made to move along a predetermined route at a speed of 3 km/h. As for the number of transmission layers, specifications were set so that the maximum number of transmission layers for each MS could be selected according to the propagation environment from a set of candidates consisting of 1, 2, 3, 4, and 8 layers. Here, it was assumed that throughputs exceeding 100 Gbps could be achieved with 4 or more layers.

4.2 Shopping Mall Scenario

1) Evaluation Results for 5G

The geometry*35 index when applying 5G and the throughput ratio of each MS are shown in **Figure 3**. For 5G, only 10 fixed BSs were used. The

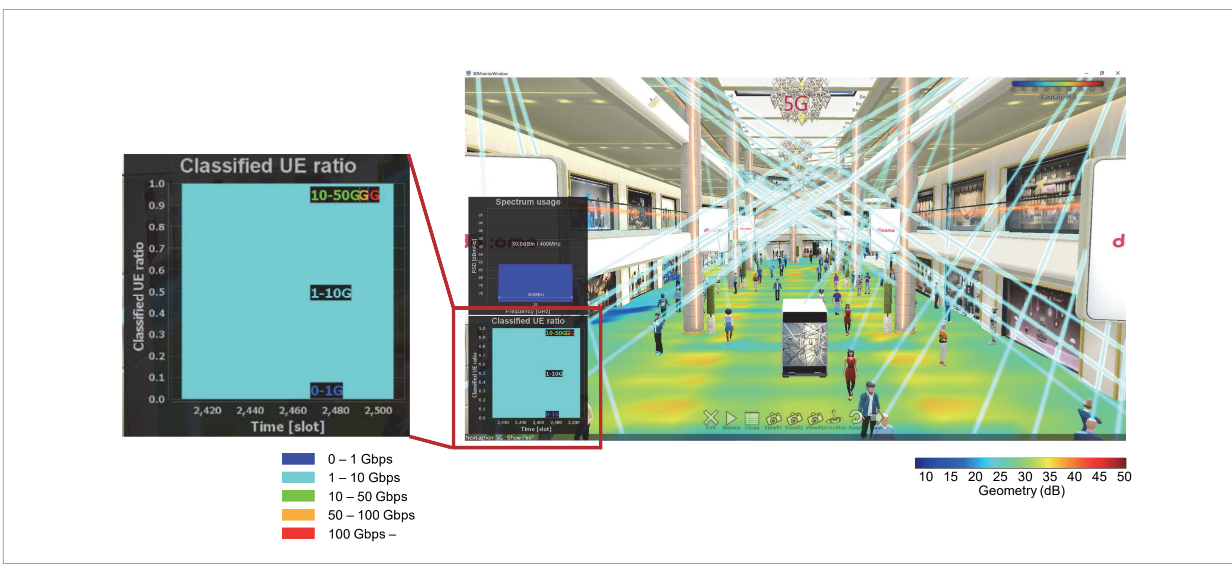


Figure 3 Evaluation results for 5G (shopping mall)

*30 Center frequency: The frequency within a frequency band at the center of the range used for communication.

*31 Sub-arrays: When generating L beams in Massive MIMO having N antenna elements, a full-array configuration generates L beams while sharing those N elements. In contrast, a sub-array generates a single beam using N/L elements. Sub-arrays are used to reduce the scale of circuitry.

*32 Flat array: An array structure featuring a two-dimensional arrangement of many elements in a Massive MIMO antenna.

*33 SU-MIMO: A technology for transmitting and multiplexing multiple signal streams by multiple antennas between a base station and terminal with one user as target.

*34 Omni-directional antenna: An antenna for which radio-wave intensity is equal in all directions. Also called a non-directional antenna.

*35 Geometry: An index indicating area quality using received power distribution, etc.

geometry is displayed on the screen in a range of 10 – 50 dB. The throughput ratios are shown by the graph at the lower left of the figure. In this graph, five different colors represent throughput values in the ranges of 0 – 1 Gbps, 1 – 10 Gbps, 10 – 50 Gbps, 50 – 100 Gbps, and 100 Gbps and greater and the horizontal axis and vertical axis represent time and throughput ratio, respectively. In addition, the colors of the lines connecting the BSs and MSs in the figure correspond to those throughput colors. It can be seen from the figure that propagation loss in 5G using the 28 GHz band was large compared with that of the Sub6 band. On the other hand, the 1st floor section could be covered in full for the most part with all MSs achieving a throughput over 1 Gbps.

2) Evaluation Results for 6G

Next, **Figure 4** shows the results of increasing the number of BS elements assuming 6G and changing the center frequency to 100 GHz while keeping the fixed BSs at the same positions as in Fig. 3. Geometry and throughput ratios are shown in the figure the same as in Fig. 3. Since the bandwidth in this case is 8,000 MHz, the power spectral density*36 dropped significantly, but since the number of elements could be greatly increased compared with 5G even for a relatively small flat array given its relationship with wavelength, high BF gain could be obtained. As a result, received power near the center of the 1st floor mall was improved by approximately 10 – 15 dB compared with 5G. Due to this wide bandwidth and high BF gain, the throughput of each MS was greatly improved. It was found that approximately 70% of the MSs could achieve a throughput over 100 Gbps while approximately 30% could achieve a throughput over

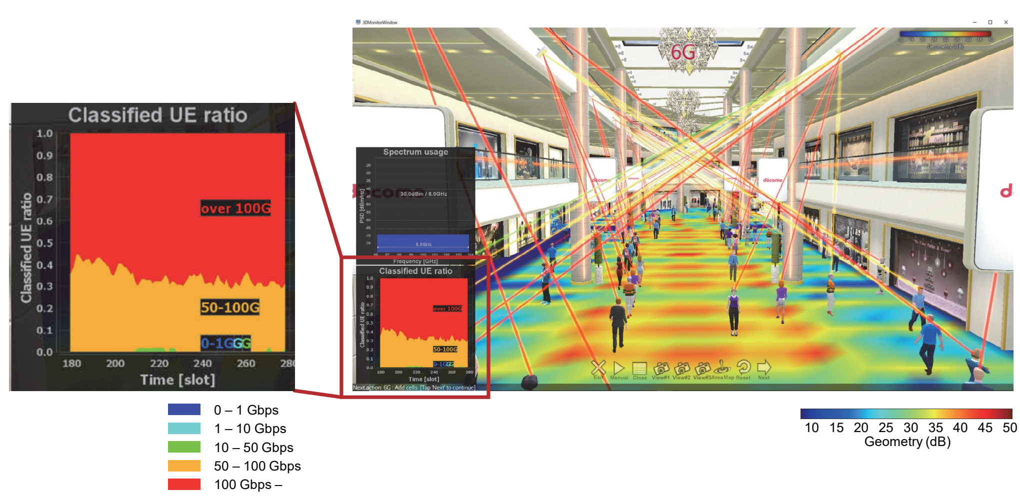


Figure 4 Evaluation results for 6G (shopping mall)

*36 Power spectral density: Power per unit frequency (1 Hz).

50 Gbps. In particular, since a single BS could generate multiple beams toward a MS, it can be seen from the colors of the straight lines connecting BSs and MSs that MSs in a Line-Of-Sight (LOS) environment and relatively close to a BS could achieve a throughput over 100 Gbps (4 or 8 layers). On the other hand, given the large propagation loss and strong straight-line propagation properties of radio waves in the 100 GHz band, received power near the walls on the 1st floor deteriorated greatly by as much as 10 dB due to the effects of blocking by the walkway on the 2nd floor. Pillars, as well, produced blocking effects so that, for example, radio waves from a fixed BS installed on a wall on the right side of the mall could not arrive at some sections of the floor on the left side.

3) Evaluation Results when Adding More BSs

Figure 5 shows evaluation results when increasing the number of fixed BSs to 20 units taking the above blocking effects into account. These results show that the increase in BS density could mitigate the effects of pillar blocking and decrease the ratio of MSs with received power less than 20 dB. Based on these results, the ratio of MSs achieving a throughput over 100 Gbps could be improved up to a value of approximately 85%. However, in this figure as well, received power remained low near the 1st floor walls.

4) Evaluation Results when Adding Drone BSs and IRSs

Figure 6 shows evaluation results when adding drone BSs and IRSs taking the above conditions into account. In this evaluation, four drone BSs moved continuously at low speed near the 1st floor walls and two IRSs were attached to each of six pillars. These results show that received power could be greatly increased near the walls of the 1st floor and generally improved by about 20 dB.

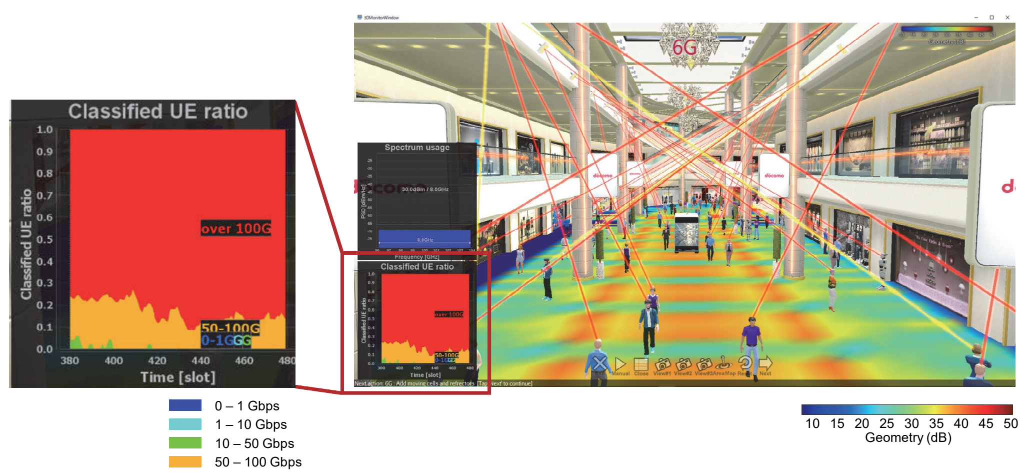


Figure 5 Evaluation results for 6G when adding more fixed BSs (shopping mall)

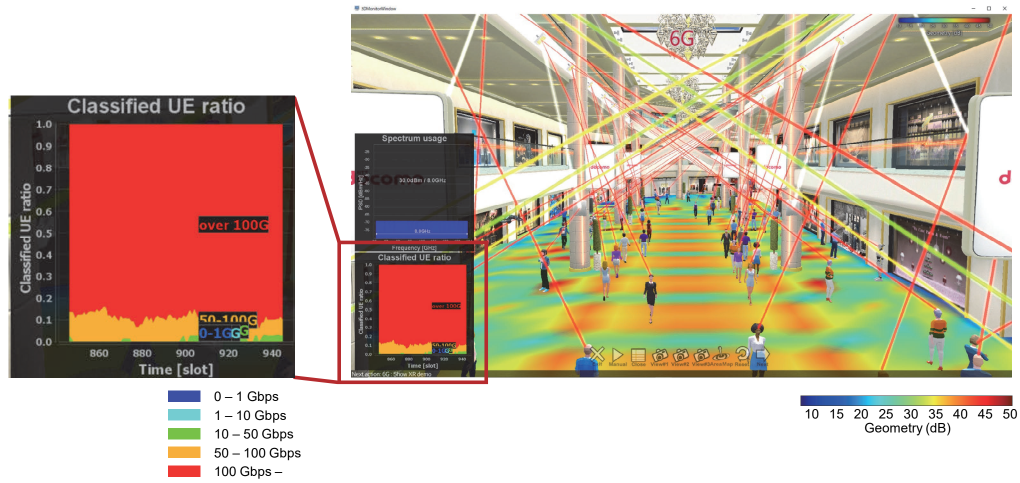


Figure 6 Evaluation results for 6G when adding IRSs and moving BSs (shopping mall)

However, since received power near the center of the mall was already high due to fixed BSs, the effects of the drone BSs and IRSs were thought to be limited to the areas near the walls. For this reason, throughput itself improved only slightly and the ratio of MSs with a throughput over 100 Gbps came to approximately 90%.

The above simulation results for the shopping mall scenario showed that throughputs over 100 Gbps could be achieved when using the 100 GHz band in 6G and that the effects of blocking could be mitigated by the use of drone BSs and IRSs.

4.3 Factory Scenario

Next, we set out to clarify geometry and throughput in the factory scenario.

1) Evaluation Results for 5G

First, evaluation results for the case of 5G are shown in **Figure 7**. For 5G, 12 fixed BSs are installed on the ceiling. In contrast to the shopping mall environment, radio waves radiate from the ceiling with little blocking in the zenith direction. As a result, received power of approximately 35 dB could be achieved in a stable manner in all areas and a throughput of 1 Gbps and higher could be achieved for all MSs.

2) Evaluation Results for 6G

Next, evaluation results when replacing 5G with 6G for the same arrangement and number of fixed BSs as 5G are shown in **Figure 8**. Similar to the shopping mall scenario, received power at locations entering a LOS state due to BF gain could be improved by more than 10 dB compared with 5G. In addition, approximately 30% of MSs could achieve a throughput over 100 Gbps while approximately 60% of MSs could achieve a throughput over 50 Gbps in combination with wideband effects. Additionally, as in the shopping mall scenario, it

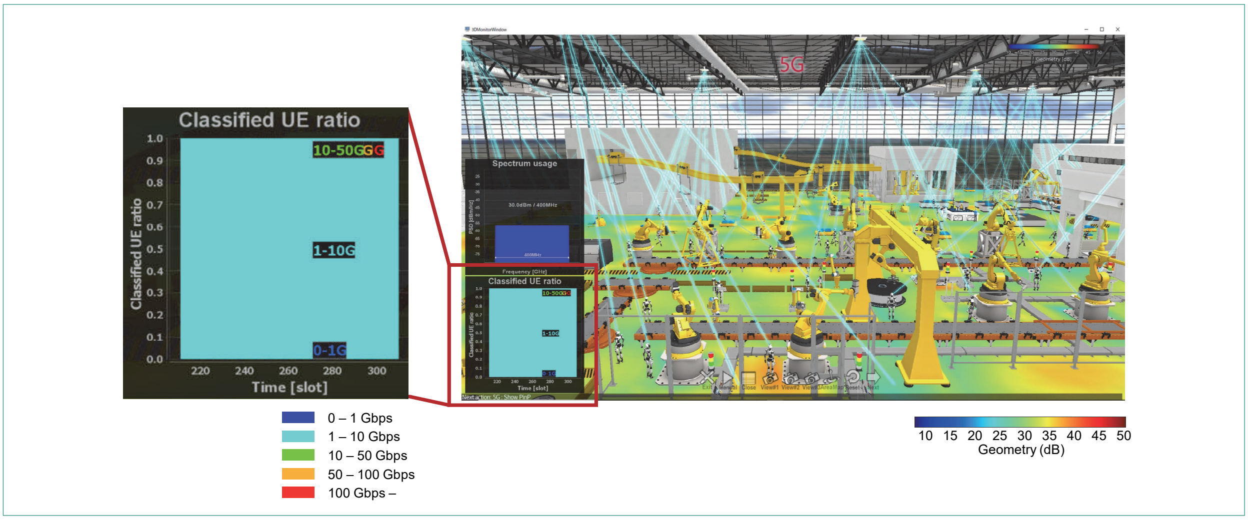


Figure 7　Evaluation results for 5G (factory)

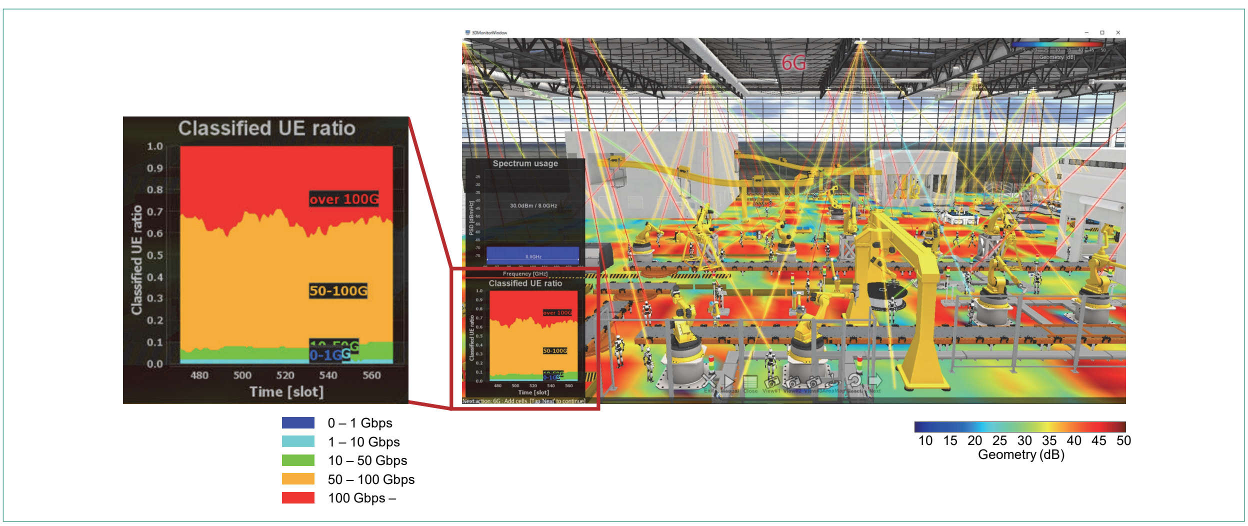


Figure 8　Evaluation results for 6G (factory)

can be seen from the colors of the straight lines connecting BSs and MSs that communications over 100 Gbps (4 or more layers) could be achieved mainly in a LOS environment. However, the factory scenario as well suffered from the effects of blocking caused by the crane or other obstructions in relation to 100 GHz propagation loss and straight-line propagation properties, and received power in some areas came to about 20 dB. As a result, the ratio of user throughput in the range of 10 – 50 Gbps was high compared with the shopping mall scenario.

3) Evaluation Results when Adding More BSs

We increased the number of fixed BSs to 25 units in response to the above blocking effects. Results are shown in **Figure 9**. This increase in the number of BSs shortened the average arrival distance and enabled BF from different fixed BSs to blocked locations. As a result, received power improved over all areas and there were almost no areas with received power less than 30 dB. This had the effect of increasing the ratio of MSs that could achieve a throughput over 100 Gbps to approximately 80% while greatly reducing the ratio of MSs having a throughput in the range of 10 – 50 Gbps.

On the other hand, received power near the structures in the background of Fig. 9 was low, which is thought to be due to the blocking of radio waves from fixed BSs above those structures.

4) Evaluation Results when Adding Moving BSs and IRSs

Taking the above blocking effects into account, we added IRSs near the above structures and near the crane in the foreground of the figure while also adding moving BSs that moved left and right in the center of the figure. Evaluation results in this case are shown in **Figure 10**. These results show that blocking effects could be mitigated and received power improved and that the ratio of MSs with a throughput over 100 Gbps could be improved to approximately 85%.

Based on the above, it was clarified for a factory scenario too that using the 100 GHz band combined with the installation of IRSs could improve received power over all areas and achieve throughputs over 100 Gbps.

5. Conclusion

This article described system-level simulations for two scenarios—a shopping mall and factory—to clarify the feasibility of extreme-high-data-rate

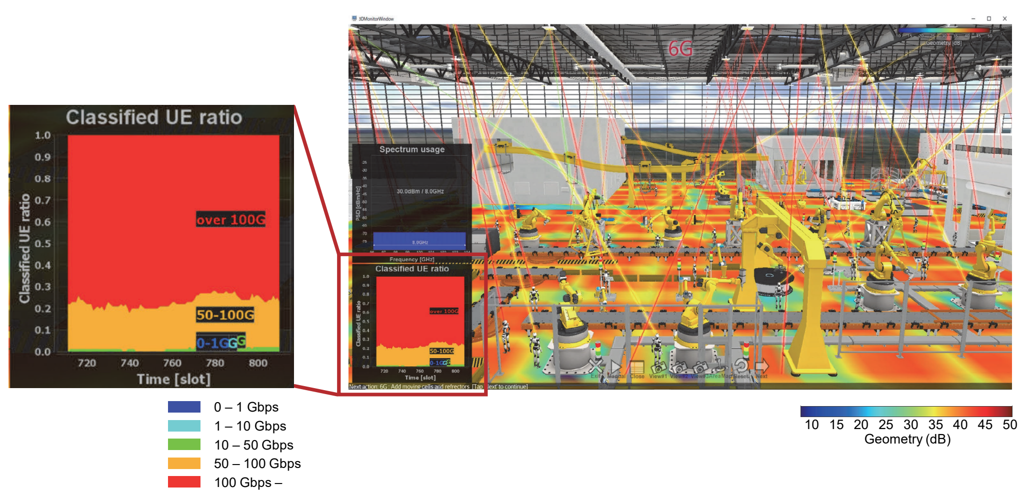


Figure 9 Evaluation results for 6G when adding more fixed BSs (factory)

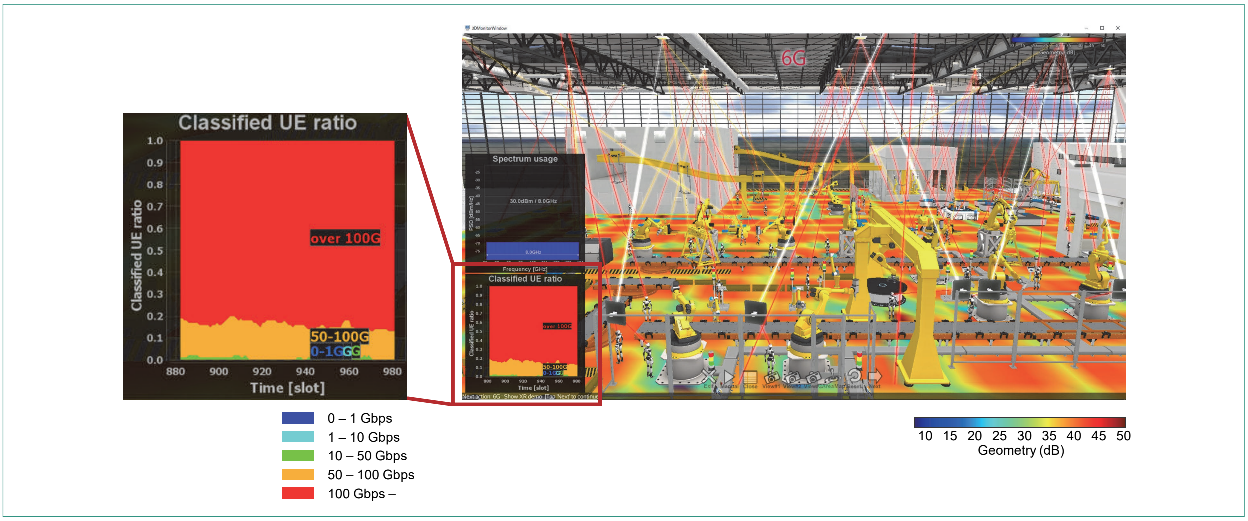


Figure 10 Evaluation results for 6G when adding IRSs and moving BSs (factory)

communications using the 100 GHz band. Going forward, there will be a need to evaluate performance for uplink communications and diverse scenarios to make 6G a reality. In future research, we seek to extend the simulator presented here so that system evaluation and performance of various technical concepts can be visualized and that use cases fitting the 6G era can be experienced in accordance with that performance level.

REFERENCES

[1] R. Alghamdi, R. Alhadrami, D. Alhothali, H. Almorad, A. Faisal, S. Helal, R. Shalabi, R. Asfour, N. Hammad, A. Shams, N. Saeed, H. Dahrouj, T. Y. Al-Naffouri and M-S. Alouini: "Intelligent Surfaces for 6G Wireless Networks: A Survey of Optimization and Performance Analysis Techniques," IEEE Access, Vol.8, Oct. 2020.

[2] P. Skrimponis, S. Dutta, M. Mezzavilla, S. Rangan, S. H. Mirfarshbafan, C. Studer and J. Buckwalter: "Power Consumption Analysis for Mobile MmWave and Sub-THz Receivers," IEEE 2020 2nd 6G Wireless Summit, Mar. 2020.

[3] NTT DOCOMO: "White Paper: 5G evolution and 6G (Version 3.0)," Feb. 2021. https://www.nttdocomo.co.jp/english/corporate/technology/whitepaper_6g/

[4] Y. Kishiyama et al.: "Trends and Target Implementations for 5G evolution & 6G," NTT DOCOMO Technical Journal, Vol.23, No.2, pp.4–13, Oct. 2021.

[5] S. Suyama, T. Okuyama, Y. Kishiyama, N. Nonaka and T. Asai: "A Study of 6G Radio Access Technologies for Achieving 100 Gbps Data Rate by Extreme Wideband Transmission Using Higher Frequency Bands," IEICE Tech. Rep., RCS2020-146, pp.66–71, Dec. 2020 (In Japanese).

[6] T. Ohgane, T. Nishimura and Y. Ogawa: "Applications of Space Division Multiplexing and Those Performance in a MIMO Channel," IEICE Trans. Commun., Vol.E88-B, No.5, May, 2005.

Technology Reports | Authentication and Authorization Technology | Security | Drones

IoT Authorization Technology for Appropriate Drone Usage

Service Innovation Department Kohsuke Yamasaki Kazuhiko Ishii

Use of drones has expanded in recent years, and legal regulations for drone flight in Japan are being created. However, in enterprises using drones, there has not been sufficient study of mechanisms for managing drone use according to regulation on worksites, which are generally at a different location than the department managing the drones. As such NTT DOCOMO has developed a system to support drone use according to regulations, using digital keys and IoT authorization technology. With the system, enterprises using drones can manage drones on the worksite using digital keys, drone use in the enterprise can be tracked from the key-use log, and unauthorized use can be prevented. This article describes the proposed technology.

1. Introduction

The term "Internet of Things" (IoT) is coming into general use in recent years, with all kinds of devices being connected to the internet. The Information and Communications White Paper from the Ministry of Internal Affairs and Communications (MIC) estimates that in 2022, approximately 35 billion devices around the world will be connected to the internet [1]. The IoT devices proliferating in this way are taking many forms, from sensor devices to automobiles, small unmanned aircraft (hereinafter refer to as "drones") and other mobile devices. As these IoT devices increase in performance, we expect that they will begin to operate more autonomously, coordinating with each other. In such a world, with IoT devices communicating and controlling each other, the authorization mechanism, which determines what is allowed, will be an important element. NTT DOCOMO has been studying various concepts for such a world, and as an IoT use case, we have applied authorization technologies to drones, which is a rapidly expanding market. This article describes some issues with enterprise use of drones, along with features of a system that we have developed.

2. Drone-use Conditions in Japan and Issues with Enterprises Using Drones

2.1 Drone Flight Rules within Japan

As of August, 2021, any flights in Japan by drones weighing 200 g or more are required to comply with rules set by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) [2]. For any flight not complying with these MLIT rules, prior application must be made and approval received from the Minister of Land, Infrastructure, Transport and Tourism, or from the regional civil aviation bureau chief. The relevant conditions are given below.

1) Airspace Requiring Permission from the Minister of Land, Infrastructure, Transport and Tourism
   (1) Airspace above and surrounding airports
   (2) Airspace at or above 150 m
   (3) Airspace above densely populated areas
2) Flights Requiring Approval from the Regional Civil Aviation Bureau Chief
   (1) Night-time flights
   (2) Flights beyond visual range
   (3) Flights within 30 m of people (3rd party) or property (3rd-party buildings, vehicles, etc.)
   (4) Flights above events
   (5) Transport of dangerous materials
   (6) Dropping of any object

A person wishing to perform a flight under any of these conditions must apply through the Drone/UAS Information Platform System (DIPS) [3], which manages such applications, and receive permission or approval. For the application, information must be submitted regarding the flight plan (date/time, location, route, etc.), the pilot (flight experience, license, etc.), the drone aircraft (design documentation giving specifications, functionality and performance of aircraft and control equipment, etc.), and a flight-safety manual. After the flight, a flight report must also be submitted through DIPS, summarizing the flight date and time, pilot, drone device, and location.

In addition to the above, regardless of aircraft weight, for flights in the vicinity of important national facilities such as the National Diet Buildings, the Prime Minister's residence, the Supreme Court, the Imperial Palace, designated official foreign residences, and nuclear power generating stations, the prefectural public safety commission must also be notified through the police station in the region where the flight will be conducted [4].

2.2 Issues with Enterprises Using Drones

As the use of drones has increased around the world, flights that violate the rules described above are becoming a problem in society [5]. Flights violating these rules can result in serious accidents, so enterprises using drones are also required to operate drones according to the rules. Conducting a flight without following the application procedure is also illegal, which is a risk in itself. However, enterprise departments managing drones generally do not have a system able to confirm that this application process has been completed when the drone flight is conducted at the worksite. This is one issue for appropriate on-site operation of drones by enterprises.

3. Summary of System Requirements

We now describe the usual configuration for use of drones and roles for drone operation within enterprises.

To create system requirements, we conducted hearings with several enterprises that use drones and summarized the results.

3.1 General Configuration When Using Drones

Except for hobby devices and some industrial devices that are highly customizable, most drones are used in the configuration shown in **Figure 1** (based on a survey of products from the top three manufactures with the largest global market share: DJI, Parrot, and 3D Robotics [6]).

Normally, the pilot flies the drone using a transmitter, which is called the controller, connected by a cable to a mobile terminal such as a smartphone or tablet. The transmitter is equipped with a controller stick, which the pilot uses to control the drone. A dedicated radio signal is used for communication between the transmitter and the drone. A drone flight application is installed on the mobile terminal, which provides functions to control the drone, and also to display video and other sensor data obtained by the drone.

3.2 Division of Roles for Drone Operation within Enterprises

Within an enterprise, drone operation is divided into the following three roles (**Table 1**).

- Drone management
- Project leader
- Pilot

The drone management department handles all drone aircraft and manages all drone use within the enterprise. The project leader is responsible for work involving a drone (such as agriculture, public works, construction, logistics, inspections, or

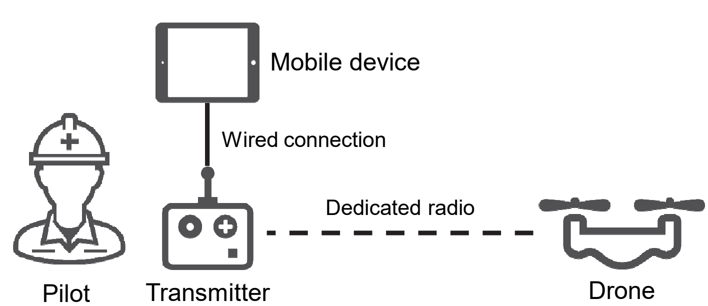


Figure 1　General configuration for drone use

measurement), studying flight plans with consideration of the work plan, and preparing applications to administrative agencies as necessary. Finally, the pilot is assigned by the project leader and performs the actual operation of the drone at the worksite. Note that the work of drone management, project leader and pilot is expected to be done at different locations.

3.3 System Requirements

We held hearings with enterprises using drones and summarized requirements for a system. These requirements are summarized in **Table 2**.

Requirements (1) to (4) are regarding flights

Table 1 Division of roles for drone operation

Staff	Location	Role
Drone management department	Head office	Handles drone hardware and usage within the company
Project leader	Office	Studies work plans and flight plans, making applications to administrative agencies as necessary
Pilot	Worksite	Operates drones at worksites

Table 2 System requirements for drone operation

Requirement (1)	Use of the drone must only be permitted if the operating pilot is the same as the pilot specified in the request. No other pilot is permitted.
Requirement (2)	Only the drone specified in the request can be used for a flight. No other drone is permitted.
Requirement (3)	Use of the drone is only permitted if the start of flight operation is within the time period specified in the request. Start of operation is not permitted at any other time. The time of completion must also be recorded, in a manner that is not easily overwritten.
Requirement (4)	Drone flight is only permitted at the location specified in the request. Flight at any other location is not permitted.
Requirement (5)	Results of the flight must be provided such that they can be checked by the drone management department and project leader.
Requirement (6)	Worksites where flight work is done must consider that connection to the internet may not be possible.
Requirement (7)	General configurations for drone use must be supported.

based on application to administrative authorities as described above. However, for requirement (3), terminating operation of a drone during flight at the planned end-of-flight time would result in safety issues, so only the flight start time is checked and end-of-flight time is recorded in the log. For requirement (5), we envision a mechanism whereby the drone management department is able to check that all company flights are conducted according to the corresponding application. For requirement (6), we learned from our hearings that in many cases, connection to the internet is not possible in the environments where drones are used. As such, it was necessary to implement a mechanism able to check that flights are conducted according to the application, even when a connection to the internet is not available. Finally, requirement (7) was included for drones already in use by enterprises, which are general-purpose drones on the market and not highly-customizable, specialized drones.

4. Prototype System Overview

Our prototype system uses a mechanism that issues a digital key with embedded conditions based on the application made to the administrative authority, and this key permits operation of the drone only under conditions consistent with the requirements, by the pilot holding the key. The system architecture is shown in **Figure 2**.

The functions of each entity*1 are described

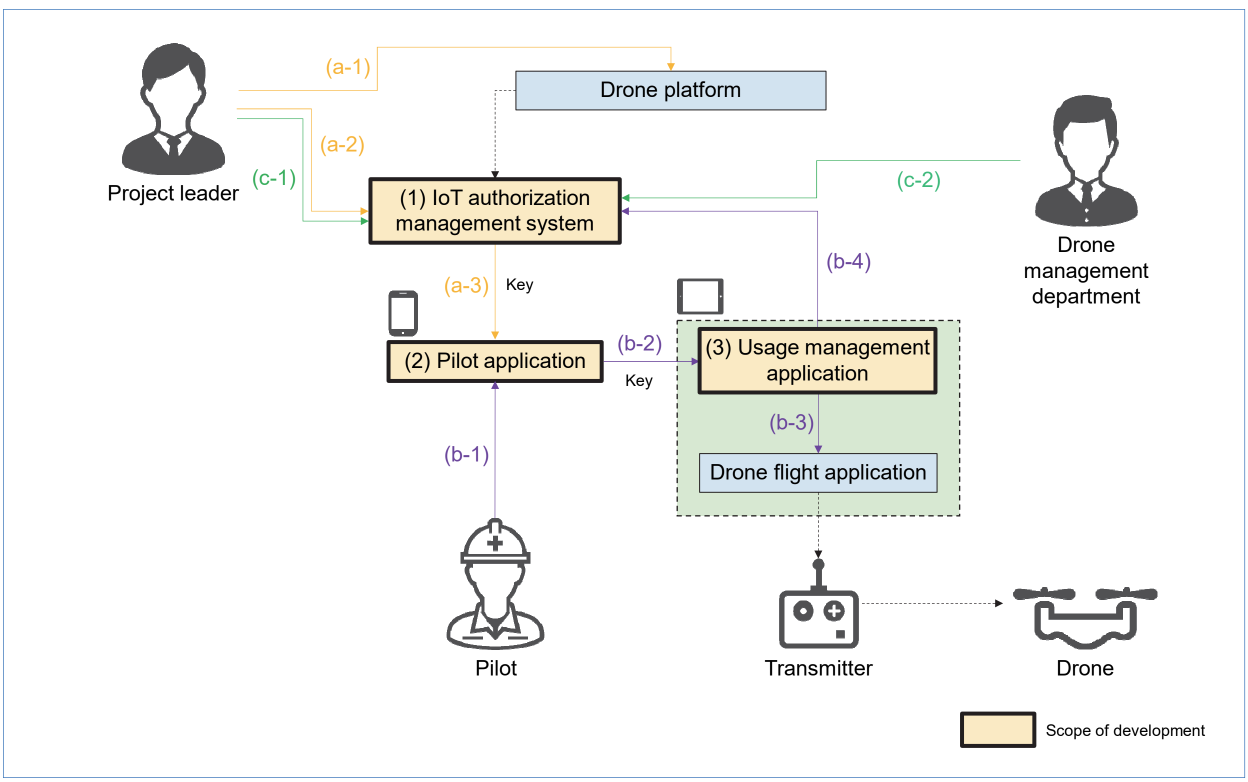


Figure 2　Prototype system configuration

*1 Entity: A constituent element providing a function in logical architecture.

below.

4.1 Functions of Each Entity

The entities we developed are described in Fig. 2 as (1) the IoT authorization management system, (2) a pilot application, and (3) a usage management application. In addition to these, an existing system with functions to support making flight plans and submitting applications to the administrative agency (hereinafter referred to as a "drone platform") is also used, and in this case we used NTT DOCOMO's docomo sky [7] application. Existing applications associated with the drone are also used for the drone flight application, and we used the DJI GO 4 application [8] from DJI with our prototype. For the drone, we used the MAVIC2Pro [9] from DJI.

(1) IoT authorization management system

The IoT authorization management system is a server application. The system is able to retrieve flight-plan data created with the drone platform through an Application Programming Interface (API)*2. It also has functions to issue, modify and add-to digital keys based on the flight plan created on the drone platform. The digital key is a document that specifies pilot, drone and time conditions and is signed with the digital signature*3 of the IoT authorization management system, to detect any falsification. When a key is issued, the IoT authorization management system transmits the key to the pilot application on the smartphone of the specified pilot. There is also a function that collects a usage log after completion of each drone flight and has a flight-list screen that highlights any flights that have exceeded the planned flight time. The system is designed to prevent unplanned flight activity using a mechanism that detects such activity. The system can also provide signed documents verifying flight logs.

(2) Pilot application

The pilot application is installed in the pilot's smartphone. The current NTT DOCOMO prototype was developed for Android 8.0. The pilot application receives a key from the IoT authorization management system, and has a function to provide the key to the mobile device connected to the transmitter. We used Near Field Communication (NFC)*4 for communication between the smartphone and the mobile device. We also use biometric authentication to verify the identity of the pilot.

(3) Usage management application

The usage management application is installed on the mobile device connected to the transmitter. The current NTT DOCOMO prototype was developed as an Android 8.0 tablet application on the same mobile device as the drone flight application. The application verifies the key received from the pilot application and checks whether the conditions are met. If a key has not been received, if the key verification fails, or if the conditions are not met, operation of the drone will be restricted, even if the drone flight application is launched. When the drone flight application is enabled, the times it is launched and terminated are recorded in the background, and this usage log is sent to the IoT authorization management system.

---

*2 API: An interface specification for different software to mutually connect.

*3 Digital signatures: A mechanism used to prevent forgery or falsification of digital data.

*4 NFC: A short-range wireless communication technology used by FeliCa and other systems.

4.2 Procedures for Usage

The steps shown in Fig. 2 are categorized into three phases: Pre-flight (a-1 to a-3), In-flight (b-1 to b-4), and Post-flight (c-1 to c-2). The procedures in each of these phases are described below.

1) Pre-flight (a-1 to a-3)

- a-1 The project leader studies the work plan and creates a flight plan using the drone platform. The leader then applies to the administrative agency and receives approval or authorization.
- a-2 The project leader logs in to the IoT authorization management system through a browser, selects the flight plan created on the drone platform, and issues a key.
- a-3 The IoT authorization management system extracts the pilot, drone and flight time from the flight plan, creates a key, and sends the key to the pilot application on the pilot's smartphone.

2) In-flight (b-1 to b-4)

- b-1 The pilot launches the pilot application and performs biometric authentication.
- b-2 The pilot selects the relevant key and sends it by NFC to the usage management application.
- b-3 The usage management application verifies the key it has received, checks the conditions, and if there are no problems, it enables the drone flight application.
- b-4 The usage management application records the times when the drone flight application was launched and terminated, and sends a log containing that information to the IoT authorization management system.

3) Post-flight (c-1 to c-2)

- c-1 The project leader obtains the data for the flight report and creates the flight report.
- c-2 Staff at the drone management department check flight status on the company flight list screen, and if a flight exceeds the permitted flight time, they can issue warnings, conduct a hearing or consider other measures to prevent recurrence.

4.3 Authentication and Key Verification between Devices

The authors' authorization concept for use of IoT is shown in **Figure 3**. A digital certificate*5 is assigned to each IoT related device, which devices use to authenticate each other. This creates an environment where permissions information can be exchanged, and a way to enforce detailed permissions requirements. In the current drone-service application, the IoT devices include smartphones, mobile devices and drones, and permissions information refers to the keys, which incorporate the permission conditions.

We now describe the implementation of our system, assuming the concepts just described (**Figure 4**). In our system, the smartphone and mobile device perform one-way authentication. The smartphone sends a key together with smartphone authentication information to the mobile device, within 60 seconds of creating the data. The usage management application on the mobile terminal then verifies the smartphone authentication information to perform the one-way authentication.

The usage management application then verifies the key by verifying the digital signature attached

*5 Digital certificate: A mechanism for preventing impersonation, issued by a trusted certificate authority.

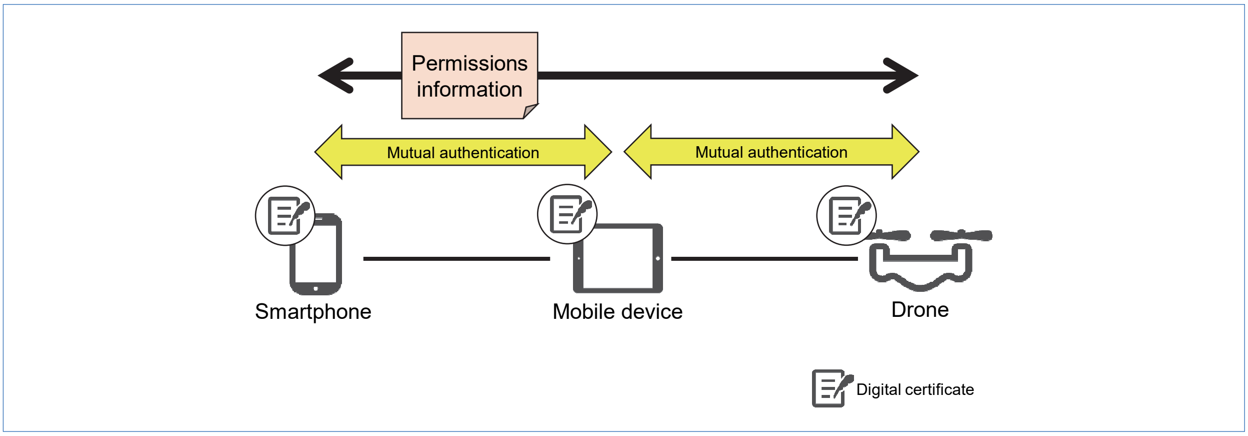


Figure 3　Authorization concept for IoT use

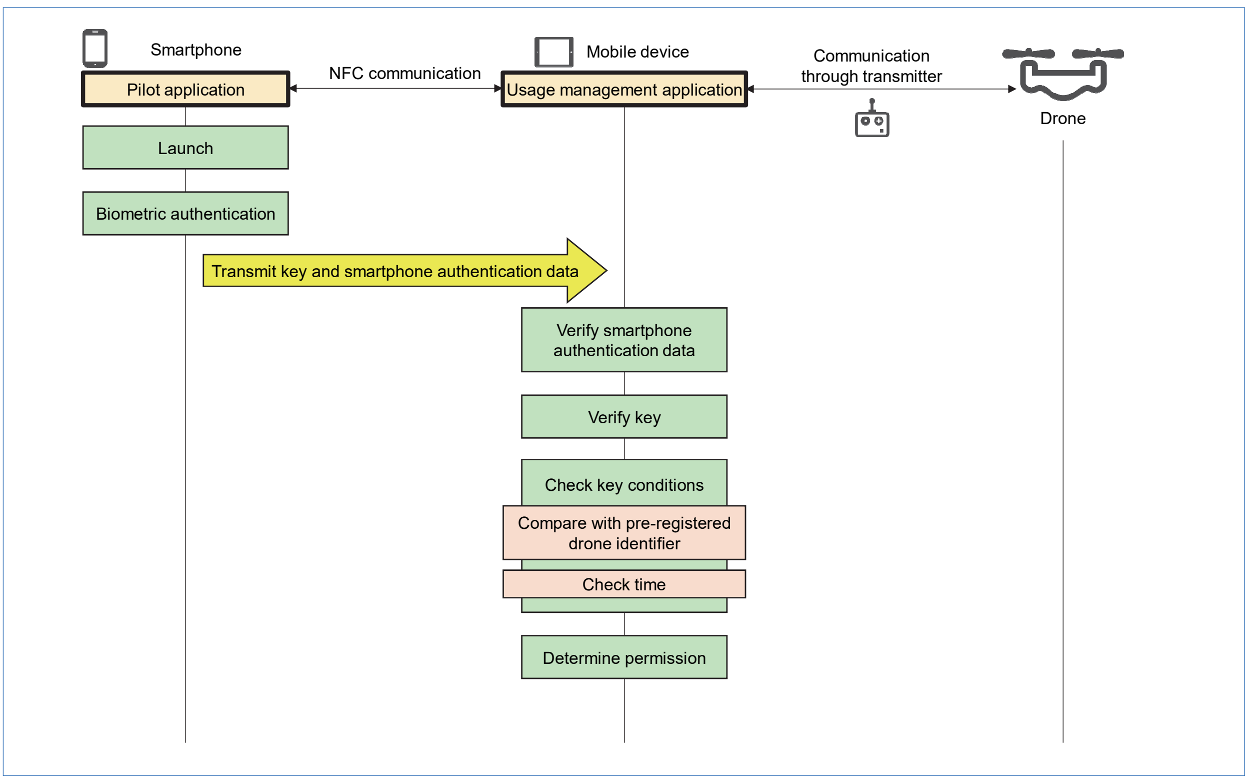


Figure 4　Authentication between devices and key verification

to it when the IoT authorization management system issued the key. To check the conditions in the key, a drone identifier that was previously registered in the usage management application is used, rather than obtaining authentication information from the drone itself. It also compares the current time in the mobile device with that stipulated in the conditions to decide whether authorization is granted. It is important to check that the pilot stipulated in the key conditions is the same as the pilot accessing the system. To ensure this, the IoT authorization management system manages pre-registered information associating the pilot, smartphone and pilot application, and sends the key to the smartphone of the pilot stipulated in the key conditions. When pilots launch the pilot application, their identity is verified biometrically before presenting the received key to the usage management application, so as a result, only the pilot stipulated in the key conditions can use the key in the usage management application. For this reason, there is no need to check the pilot identity within the usage management application.

4.4 Implementation of System Requirements in the System

We now describe how the above system requirements are implemented in the system. Requirements (1) to (3) are embedded in the key issued in a-3, and the usage management application checks the conditions in the key it receives in b-2, giving authorization based on the result, which enforces these requirements. For requirement (3), the flight-end time is recorded in the usage log, which the pilot cannot edit. The pilot application maintains this log and sends it to the IoT authorization management system. Requirement (4) was not implemented in the current prototype. Requirement (5) is implemented with the usage log obtained in b-4, and by showing the drone usage state on a management screen, where it can be checked. Requirement (6) is implemented by enabling the pilot to complete operation steps b-1 to b-3 at worksites, even when an internet connection is not available. To implement requirement (7), the usage management application monitors the drone flight application in the background, and is able to limit use of this application. The drone flight application is not fixed, and the system can be used with various drone flight applications.

5. Discussion

We discuss several issues arising in development of the system below.

1) Drone Aircraft Authentication

With the current prototype, it is not possible to authenticate the drone aircraft itself. This is because current general-purpose drones do not have an authentication function. The MAVIC2 PRO [9] from DJI, which we used, has a Mobile Software Development Kit (SDK)*6 [10] that enables a serial number to be retrieved from the drone as an identifier. However, we were unable to run an application using the Mobile SDK on the mobile device where the drone flight application was installed. Thus, for the prototype, we implemented a function to record the drone identifier on the mobile device beforehand. Authentication and exchange of permissions information is done between the smartphone and the mobile device and is not done with the drone. Ultimately, it would be better to authenticate the drone itself, and compare against the conditions in the key.

2) Authentication between Smartphone and Mobile Device

For our prototype, we implemented one-way authentication from the mobile device to the smartphone using NFC communication. Two-way authentication would be preferable, but due to development-time

*6 Mobile SDK: Software required for developing applications on mobile devices such as smartphones.

constraints, only one-way authentication was implemented in our prototype. This leaves open the possibility of attack, stealing a key by impersonating the mobile device. To prevent this, we added a time limit on the data sent by NFC from the smartphone, separate from the time limit of the conditions described in the key, so that the authorization cannot be enacted if the data sent by NFC is leaked. More specifically, if a person obtaining the leaked data attempts to use it, the usage management application will verify the time limit on the transmitted data and prevent authorization. To achieve this, the time limit on the transmitted data is shorter than the limit stipulated in the key conditions.

3) Preventing Over-use

The prototype system is not able to prevent flights from exceeding the authorized time. As such, to discourage operation beyond the authorized time, we highlight flights that exceed the planned time on the flight-list screen, so that the drone management department will be alerted. It will still be necessary to verify whether this mechanism works to reduce use outside of the authorized time in actual operation.

4) Internal Unauthorized Use

With the prototype system, it is still possible to use a completely new mobile device to fly a drone without receiving authorization, but even in such a case, an inconsistency would be created in the usage logs, making it possible to detect the unauthorized use. However, if a legitimate pilot enables the drone flight application and then gives it to another unauthorized pilot, that pilot will be able to operate the drone. The current prototype system is not able to prevent such unauthorized use. One possible way to counter this would be to use the user-facing camera on the mobile device to photograph the pilot during operation, so the pilot could be verified afterward. In the future we will investigate how strictly to enforce restrictions in relation to cost, among other issues, as we continue to conduct tests with enterprises that are actually using drones.

6. Conclusion

This article describes a system created to support operation of drones according to authorized flight conditions. The system was designed based on requirements compiled from hearings held with enterprises that are actually using drones. The system embeds conditions in a digital key, based on an application made to the applicable administrative agency. It then only permits a pilot with the key to use the drone, at the worksite and according to the approved flight plan. We studied the prototype system further, clarifying how well we can maintain security in current usage environments, and what further functionality is needed to create a more ideal system. As a next step in the future, we will conduct further tests with enterprises using drones, evaluate the utility of the system, and examine what functionality is needed in real environments. The rules regarding drones are also being updated each year, so the system design will need to enable the new rules to be followed.

In the future, we also want to apply the IoT authorization management system to use cases other than drones, to implement a mechanism by which IoT devices can collaborate autonomously, while reliably enforcing authorization with each other.

REFERENCES

[1] Ministry of Internal Affairs and Communications: "2020 Information and Communications White Paper," (Accessed May 19, 2021) (In Japanese).
https://www.soumu.go.jp/johotsusintokei/whitepaper/r02.html

[2] Ministry of Land, Infrastructure, Transport and Tourism: "Flight rules for unmanned aircraft (drones, radio-controlled, etc.)," (Accessed May 19, 2021) (In Japanese).
https://www.mlit.go.jp/koku/koku_tk10_000003.html#a

[3] Ministry of Land, Infrastructure, Transport and Tourism: "Drone information platform system," (Accessed May 19, 2021) (In Japanese).
https://www.dips.mlit.go.jp/portal/

[4] National Police Agency: "Regarding prohibition of small unmanned aircraft flights," (Accessed May 19, 2021) (In Japanese).
https://www.npa.go.jp/bureau/security/kogatamujinki/index.html

[5] Nippon Keizai Shinbun: "Exposing illegal drone flights: A record 111 in 2019," Mar. 2020 (In Japanese).
https://www.nikkei.com/article/DGXMZO57242650W0A320C2MM0000/

[6] Japan Patent Office: "FY2018 Survey Report on Technology Trends in Patent Applications: Drones," Feb. 2019 (In Japanese).
https://www.jpo.go.jp/resources/report/gidouuhoukoku/tokkyo/document/index/30_05.pdf

[7] NTT DOCOMO Co. Ltd.: "The docomo sky Drone Platform," (In Japanese).
https://www.docomosky.jp/

[8] DJI: "DJI GO 4," (In Japanese).
https://www.dji.com/jp/downloads/djiapp/dji-go-4

[9] DJI: "MAVIC 2," (In Japanese).
https://www.dji.com/jp/mavic-2?site=brandsite&from=nav

[10] DJI: "DJI DEVELOPER," (In Japanese).
https://developer.dji.com/

Technology Reports

CUPS | EPC/5GC | Traffic Characteristics

CUPS for Flexible U-Plane Processing Based on Traffic Characteristics

Core Network Development Department Yuya Miyazaki Kenzo Okuda
Kouichiro Kunitomo
DOCOMO Technology, Inc. Packet Network Division Takahiro Kaida

As 5G services become more widespread, it is imperative that EPC further address various traffic characteristics such as low latency communications and high-capacity user communications. NTT DOCOMO has introduced CUPS architecture that enables functional separation of C-Plane and U-Plane parts from SGW and PGW. This separation realizes flexible U-Plane handling based on traffic characteristics. This article provides an overview of CUPS technology and describes how facilities to process U-Plane are selected.

1. Introduction

At the initial deployment phase of 5th Generation mobile communication systems (5G), the 5G Non-Stand-Alone (NSA) architecture*1 was widely adopted to realize 5G services by connecting 5G base stations to the existing Evolved Packet Core (EPC)*2 [1]. As applications based on 5G become more widespread, the need for EPC to achieve higher speed and capacity communications, lower latency communications and simultaneous connection of many terminals than ever has become urgent [2] [3]. Specifically, it is necessary to increase the number of high-capacity gateway devices capable of processing hundreds of Gbps to several Tbps to achieve high-speed, high-capacity communications,

*1 5G NSA architecture: A 5G system configuration in which 5G radio base stations and EPCs are linked. This enables a low barrier to deployment because it allows 5G to be provided without installing 5GC (See *38).

*2 EPC: The fourth-generation IP-based core network specified by 3GPP for LTE and other access technologies.

to distribute gateway devices near base station facilities to achieve even lower latency communications, and to improve session*3 processing performance for connecting massive numbers of terminals simultaneously.

Conventional single gateway devices have both Control Plane (C-Plane)*4 functions to manage communication sessions and control communications, and User Plane (U-Plane)*5 functions to handle communications traffic [4]. Therefore, if the previously assumed balance between the number of sessions and communications capacity is disrupted, either the C-Plane or the U-Plane will have excess processing capacity. In high-speed, high-capacity communications, the C-Plane has excess processing power, and in multiple terminal simultaneous connections, the U-Plane has excess processing power because the volume of communications is small compared to the number of sessions. If the C-Plane and U-Plane can be scaled*6 independently, these issues can be resolved, and efficient facility design can be expected. In addition, low-latency communications require distributed deployment of the U-Plane function near the base station facilities to reduce propagation delay. However, in the distributed deployment of conventional devices with integrated C-Plane and U-Plane functions, the number of sessions and communication volume are unevenly distributed among the gateway devices, resulting in a decrease in the efficiency of facility utilization. Since there is no need for distributed deployment of C-Plane functions, if the C-Plane and U-Plane functions can be separated and the way they are deployed changed according to their characteristics, the loss of facility utilization efficiency related to C-Plane processing capacity could be greatly reduced.

From above background, NTT DOCOMO has been planning to deploy Control and User Plane Separation (CUPS)*7 architecture to realize the separation of C-Plane and U-Plane functions as specified in 3rd Generation Partnership Project Technical Specification (3GPP TS) 23.214. Separating the C-Plane and U-Plane functions of gateway devices with CUPS architecture makes it possible to scale the C-Plane and U-Plane independently and balance the centralized deployment of C-Plane functions with the distributed deployment of U-Plane functions, thereby enabling the deployment and development of a flexible and efficient core network*8 [5]. In addition to solving the aforementioned issues, CUPS will also enable independent equipment upgrades for C-Plane and U-Plane functions, and the adoption of U-Plane devices specialized for specific traffic characteristics.

In the user perspective, the introduction of CUPS can be expected to dramatically improve the user experience through the operation of facilities specializing in various requirements, and enable further increases in facilities and lower charges to pursue user benefits by improving the efficiency of core network facilities.

Regarding the CUPS architecture, a source of value for both operators and users, this article includes an overview of the architecture, additional control protocols, U-Plane control schemes based

*3 Session: A generic term that includes a virtual communications channel for exchanging data in the U-Plane, data exchanged in the communications channel, and metadata such as management information exchanged in the C-Plane about the communication channel.

*4 C-Plane: A series of control processes that are exchanged to establish communications and authentication.

*5 U-Plane: The process of sending and receiving user data, which are the main signals for communication.

*6 Scaling: Designing the capacity of processing resources and planning the increase or decrease of communications facilities based on the performance characteristics of devices, traffic models reflecting communication characteristics, expected number of subscribers, and spare capacity for burst traffic.

*7 CUPS: An architecture in which the C-Plane and U-Plane functions of SGW and PGW are separated and specified as separate devices.

*8 Core network: A network consisting of gateway devices, location management devices, subscriber information management devices, and mobility management functions. The core part of the network that constitutes a mobile communication system. A mobile device communicates with the core network via a radio access network consisting of radio base stations.

on traffic characteristics, and future developments toward a 5G Stand-Alone (5G SA) architecture*9.

2. CUPS

2.1 CUPS Overview

1) Concept and Architecture

The architecture of the EPC with CUPS is shown in **Figure 1**. CUPS is an architecture defined in 3GPP TS23.214 that separates the Serving GateWay (SGW)*10/Packet data network GateWay (PGW)*11 configuration of the EPC into the C-Plane and U-Plane. The CUPS architecture is designed so that there is no difference in the interface between the existing architecture and the CUPS architecture - even with CUPS architecture deployed in SGW/PGW, opposing devices such as a Mobility Management Entity (MME)*12, Policy and Charging Rules Function (PCRF)*13, evolved NodeB (eNB)*14/next generation NodeB (gNB)*15, and SGWs/PGWs of other networks such as Mobile Virtual Network Operator (MVNO)*16 and roaming*17 are not affected. For C-Plane, SGW Control plane function (SGW-C)*18/PGW Control plane function (PGW-C)*19, and for U-Plane, SGW User plane function (SGW-U)*20/PGW User plane function (PGW-U)*21 are equipped with call processing functions. By introducing CUPS, C-Plane/U-Plane capacities can be expanded individually as needed. Combined SGW-C/PGW-C*22 and Combined SGW-U/PGW-U*23 can

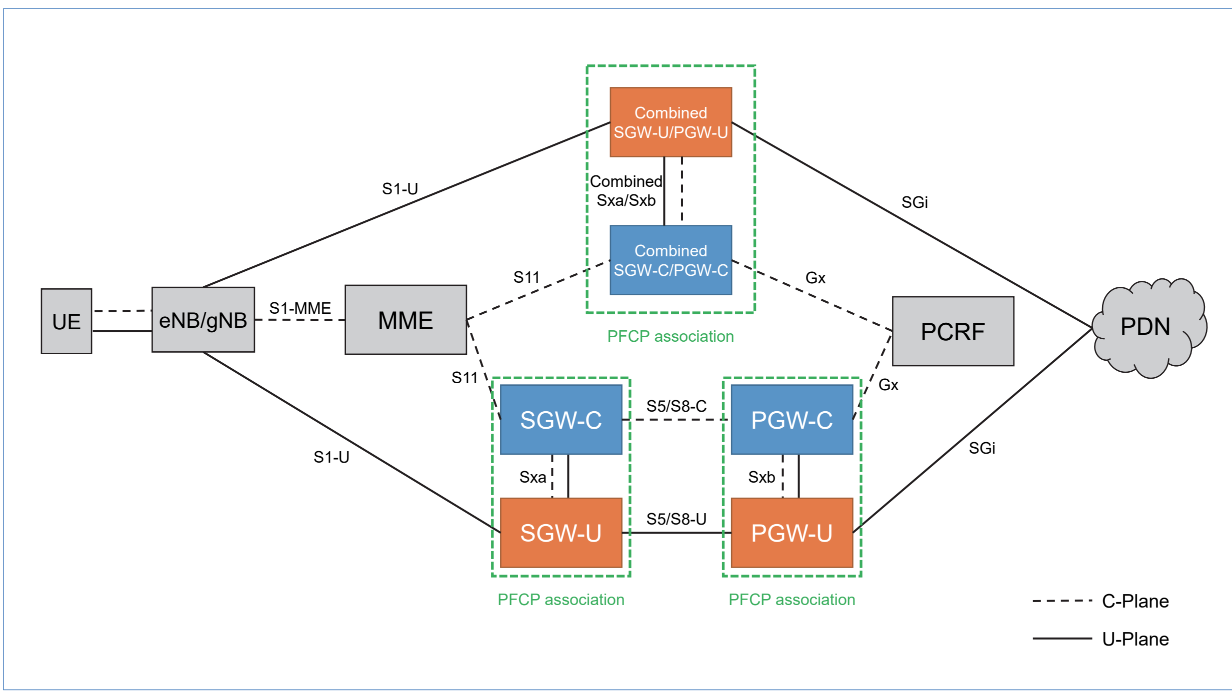


Figure 1 EPC architecture with CUPS introduced

*9 5G SA architecture: One of the system configurations of 5G. 5G is provided by 5G radio base stations and 5GCs. It is referred to as Stand-Alone configuration because it is an all-5G system that makes 5G-specific function available.

*10 SGW: The gateways (Serving GW and PDN GW) deal with the user plane. They transport the IP data traffic between the User Equipment (UE) and the external networks. The Serving GW is the point of interconnect between the radio-side and the EPC. As its name indicates, this gateway serves the UE by routing the incoming and outgoing IP packets. (source: 3GPP)

*11 PGW: The PDN GW is the point of interconnect between the EPC and the external IP networks. PDN stands for "Packet Data Network". The PDN GW routes packets to and from the PDNs (source: 3GPP).

*12 MME: The MME (Mobility Management Entity) deals with the control plane. It handles the signaling related to mobility and security for E-UTRAN access (source: 3GPP).

*13 PCRF: A logical device that manages and controls user billing policies.

*14 eNB: A radio base station that supports the LTE radio system.

*15 gNB: A radio base station that supports the 5G radio system.

handle the functions of SGW and PGW in common devices. In the standard specification, in addition to SGW/PGW, the Traffic Detection Function (TDF)*24 can also be separated into TDF-C and TDF-U, but the details are omitted in this article.

SGW-C/PGW-C have the C-Plane interfaces (S5/S8-C, S11, Gx, etc.) to handle the C-Plane and SGW-U/PGW-U have the U-Plane interfaces (S1-U, S5/S8-U, SGi, etc.) to handle the U-Plane. Interfaces that handle both C-Plane and U-Plane are separated into different interfaces for C-Plane and U-Plane. For example, for S5, SGW-C/PGW-C and SGW-U/PGWU have S5-C/S5-U, respectively. Between SGW-C/PGW-C and SGWU/PGW-U there are Sx interfaces as defined in 3GPP TS29.244 [6]. The interface between SGW-C and SGW-U is called Sxa, between PGW-C and PGW-U is called Sxb, and between Combined SGWC/PGW-C and Combined SGW-U/PGW-U is called Combined Sxa/Sxb.

For the Sx interface, the Packet Forwarding Control Protocol (PFCP)*25 specified in 3GPP TS29.244 is used. PFCP is described in detail later. It is also possible to send U-Plane over the Sx interface with GPRS Tunneling Protocol for User Plane (GTP-u)*26.

2) Standard Call Processing in CUPS

The call processing sequence for SGW/PGW with CUPS architecture is designed so that it does not affect existing architecture. The only difference between CUPS and non-CUPS architectures is more signaling between SGW-C/PGW-C and SGW-U/PGW-U. An example of LTE Attach*27 is shown in **Figure 2**. The MME that receives the Attach request selects the SGW-C/PGW-C respectively, and sends a session establishment request to the selected SGW-C (Fig. 2 (1)). SGW-C selects the SGW-U for establishing the session (Fig. 2 (2)). The SGW-C sends the PFCP session establishment request to the SGW-U for establishing the session (Fig. 2 (3)). The SGW-C then sends a session establishment request to the PGW-C (Fig. 2 (4)). Upon receiving this, PGW-C obtains the billing policy and other information for the session to be established from PCRF as well as non-CUPS architecture. The PGW-C that establishes the session next selects the PGW-U and sends the PFCP session establishment request to establish the session (Fig. 2 (5)). SGW-C sends the U-Plane path information about PGW-U obtained from the response to the session establishment request from PGW-C to SGW-U, and establishes the U-Plane path between SGW-U and PGW-U by changing the PFCP session (Fig. 2 (6)). After the MME completes the configuration of the eNB side, the SGW-C receives the eNB U-Plane path information. SGW-C notifies SGW-U of the eNB-side U-Plane path information, and SGW-U establishes the U-Plane path with the eNB side (Fig. 2 (7)).

2.2 Control between CUs

1) PFCP Overview

PFCP is a protocol used by SGW-C/PGW-C that controls packet forwarding between SGW-U/PGW-U with a set of rules. The protocol enables SGW-C/PGW-C to assign a set of rules to SGW-U/SGW-U on a per session basis.

PFCP signals are classified into two types: Node

*16 **MVNO:** A mobile communication service provider that does not own the radio access network infrastructure over which it provides services to customers.

*17 **Roaming:** A mechanism that enables mobile subscribers to receive services from visited networks when traveling outside the geographical coverage area of their home networks.

*18 **SGW-C:** The functional section that performs SGW C-Plane processing in CUPS architecture.

*19 **PGW-C:** The functional section that performs PGW C-Plane processing in CUPS architecture.

*20 **SGW-U:** The functional section that performs SGW U-Plane processing in CUPS architecture.

*21 **PGW-U:** The functional section that performs PGW U-Plane processing in CUPS architecture.

*22 **Combined SGW-C/PGW-C:** A single C-Plane processing functional section that combines SGW-C and PGW-C and combines the functions of both.

*23 **Combined SGW-U/PGW-U:** A single U-Plane processing functional section that combines SGW-U and PGW-U and combines the functions of both.

*24 **TDF:** A functional section that identifies traffic, detects applications, and notifies the PCRF. It is not implemented by NTT DOCOMO.

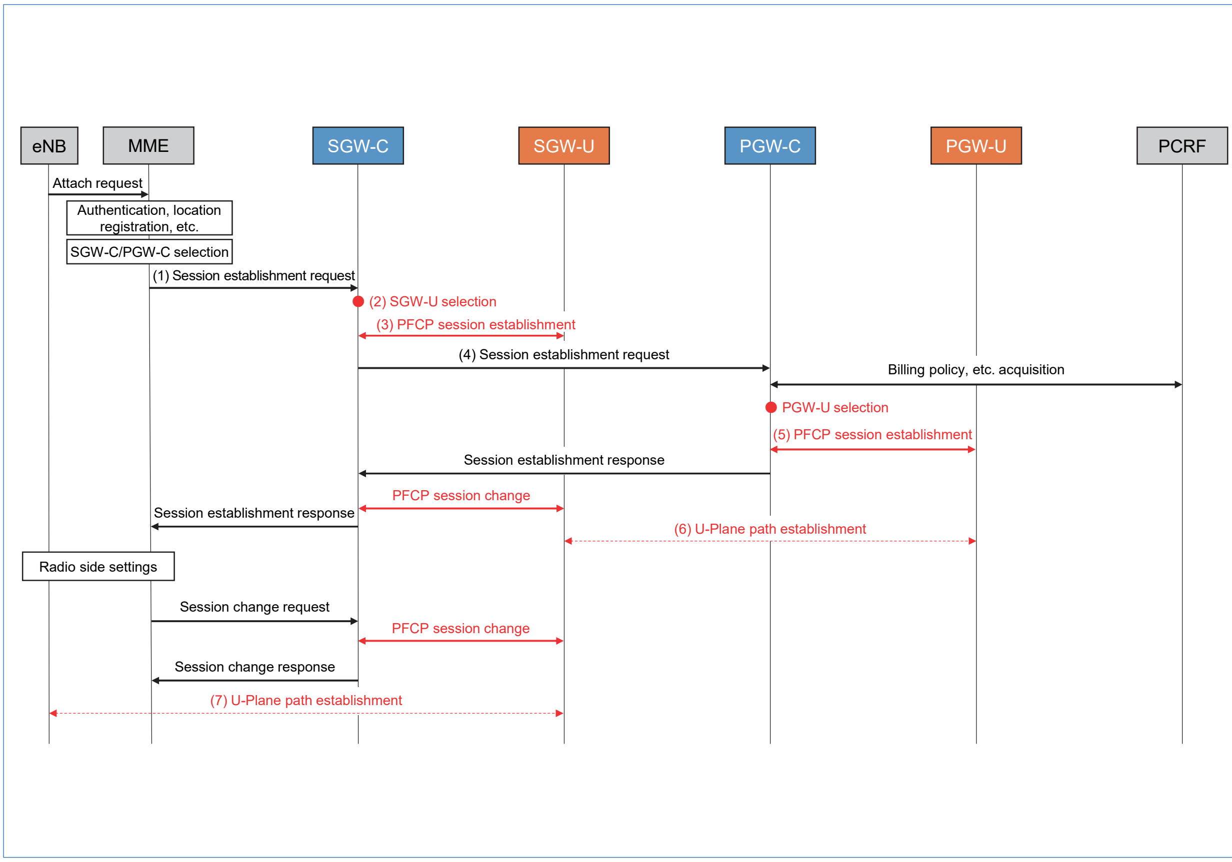


Figure 2 Attach sequence

Related Procedures/Messages that specify inter-device control and Session Related Procedures/Messages that specify PFCP session control. The SGW-C/PGW-C and SGW-U/PGW-U pair is called a PFCP association (Fig. 1). SGW-C/PGW-C can only control SGW-U/PGW-U in a PFCP association.

2) Packet Processing Model

In PFCP, multiple rules are combined to control packet processing. The sets of rules used in PFCP are shown in **Table 1**. There are five types of rules: Packet Detection Rule (PDR), Forwarding Action Rule (FAR), Buffering Action Rule (BAR), Quality of Service (QoS)*28 Enforcement Rule (QER), and Usage Reporting Rule (URR). In PFCP, packet processing is achieved by combining various rules around PDR. In PDR, the receiving interface and 5-tuple information*29 to be monitored are defined. When a packet is received, the SGW-U/PGW-U judges whether the 5-tuple information associated with the packet meets the PDR conditions. If the conditions are met, packet processing is performed according to the specified set of rules. For PDR,

*25 PFCP: C-Plane protocol used at the Sx reference point, where SGW-C/PGW-C instructs SGW-U/PGW-U on the packet control method using PFCP.

*26 GTP-u: A tunneling protocol used by radio base stations and devices in the core network to transmit user data.

*27 Attach: The processing of registering a mobile UE with a network when UE power is turned on, or the state of being registered.

*28 QoS: A technology for properly managing communications quality on a network by marking packets and giving them priority in processing, such as giving priority over data transfer to avoid interruptions in voice calls, etc.

*29 5-tuple information: A generic term for five pieces of information stored in the IP header and Transmission Control Protocol (TCP)/User Datagram Protocol (UDP) header: the destination IP address, destination port number, source IP address, source port number, and protocol number.

Table 1 PFCP rule groups

Rule name	Role
PDR	Specifies judgment conditions for received packets and the set of rules to be applied to the packets.
FAR	Specifies whether or not the packet is to be forwarded by discarding or buffering it, in addition to operations related to forwarding such as the tunneling header information to be added to the packet and the interface to be used.
BAR	Specifies the maximum number of packets to be retained and the dwell time between the arrival of a downlink packet and the notification to the SGW-C, etc. for buffering.
QER	Specifies the QoS of forwarded packets, such as the allowed bandwidth and the DSCP value assignment.
URR	Specifies how to count the packets detected by the associated PDR and when to notify the C-Plane device of the count status.

settings are made individually, such as PDR for detecting downlink packets and PDR for detecting uplink packets. Also, it is possible to change the rules to be applied for each destination IP address or protocol, for example, it is possible to specify that packets related to Dynamic Host Configuration Protocol for IP version 6 (DHCPv6)*30 are forwarded from PGW-U to PGW-C. Except PDR, it is possible to associate each rule with multiple PDRs. For example, by associating one QER with multiple PDRs, it is possible to have the same Differentiated Services Code Point (DSCP)*31 value assigned by SGW-U/PGW-U.

Now we describe the actual flow of applying rules. At first, SGW-C/PGW-C informs the multiple PDRs and the respective rules associated with them to SGW-U/PGW-U. Second, the SGW-U/PGW-U judges the group of PDRs notified when packets are received in order, and searches for matching PDRs. If a matching PDR is found, the received packet is processed based on the FAR associated with the PDR. If the content of the FAR is buffering, buffering is performed based on BAR. When forwarding packets, DSCP marking and bandwidth controls are performed based on QER. Lastly, the packet count and volume count are performed using the methods specified in URR, and if it is necessary to notify the SGW-C/PGW-C, the count information and other information is notified. In this way, PFCP rules are combined to perform packet processing in SGW-U/PGW-U.

3) U-Plane Device Management

SGW-C/PGW-C establish PFCP associations for SGW-U/PGW-U and manage these associations before actual U-Plane processing. The SGW-U/PGW-U that is the candidate for selection at the time of session establishment is the device that established the PFCP association. When establishing a PFCP association, it is possible to choose whether to share the optional functions of the standard specification between SGW-C/PGW-C and SGW-U/PGW-U. SGW-C/PGW-C and SGW-U/PGW-U exchange their

*30 DHCPv6: A protocol for distributing DNS server information, address information, and other information necessary for connecting to a network using IPv6.

*31 DSCP: A value that indicates the priority of a packet when controlling the QoS priority of IP packets. It is represented by the first six bits of the Type of Service in the IP header. 64 levels of priority can be specified.

capabilities in Node Related Procedures/Messages and decide which one will support the functions depending on how optional functions are handled.

SGW-C/PGWC and SGW-U/PGW-U that establish a PFCP association send heartbeat packets*32 periodically to each other for alive monitoring and confirming restart time. Heartbeat packets can be sent and received in both directions between these devices.

3. U-Plane Control Based on Traffic Characteristics

3.1 GW Selection after CUPS Introduction

1) CUPS GW Selection Method

For EPC, the MME selects the SGW/PGW using Domain Name System (DNS)*33. The MME selects an appropriate SGW/PGW using DNS based on key information such as Access Point Name (APN)*34 and location of the terminal. Before the introduction of CUP architecture, the MME determines the SGW/PGW by considering both the C-Plane perspective, such as number of sessions, and the U-Plane perspective, such as expected traffic volume and physical distance. In contrast, after the introduction of CUPS, the MME and SGW-C/PGW-C will share the responsibility of selecting U-Plane path. The MME will select SGW-C/PGW-C, and SGW-C/PGW-C will select SGW-U/PGW-U. Two types of SGW-U/PGW-U selection methods are specified in the standard specification. The first is that SGW-C/PGW-C selects SGW-U/PGW-U alone. The second is that SGW-C/PGW-C selects SGW-U/PGW-U cooperating with DNS. In the standard specification, the second method is optional, so this article describes the first method.

The three main parameters that can be used for SGW-U/PGW-U selection as specified in the standard specification are as follows:

(1) C-Plane information received in the session establishment request, including APN and terminal location information.

(2) Static information about the performance and functions of SGW-U/PGW-U.

(3) Dynamic information such as the load status of SGW-U/PGW-U.

Basically, the selection of SGW-U/PGW-U is the same as selection by a conventional MME, using APN and location information, but (3) leads to the realization of functions that were difficult to achieve with the selection by the MME. An example of (3) is notification of the current load status as a parameter for load balancing, although even in the past there have been specifications for notifying the MME of the load status from the SGW/PGW and utilizing it. However, since the necessity of utilizing the notified SGW/PGW load information depends on the implementation status of the MME, there was a possibility that the information would not be properly utilized by roaming and MVNOs. After the introduction of CUPS, load balancing for SGW-U/PGW-U and selection according to traffic is done by SGW-C/PGW-C on the operator's network so devices can be selected based on the load information.

*32 Heartbeat packet: A packet for survival confirmation or its survival response sent to monitor the life and death of an opposing device. Timestamps and other information are sent and received in PFCP heartbeats.

*33 DNS: A function that resolves domain names and IP addresses on a network. In the core network, it is used for service discovery for gateway devices, etc.

*34 APN: An identifier that specifies the connection destination for the UE (See *36), used by the UE as an identifier to specify the PDN to connect to when requesting a connection to the core network.

2) Device Deployment after CUPS Installation

A **Figure 3** shows an example of device deployment after CUPS installation. After the introduction of CUPS, SGW-C/PGW-C and SGW-U/PGW-U can be deployed independently. In the standard specification, SGW-C/PGW-C and SGW-U/PGW-U can be connected N to N. For example, SGW-C/PGW-C can be redundantly distributed in consideration of disasters and congestion, so that even if one SGW-C/PGW-C is out of order, SGW-U/PGW-U can be selected from another SGW-C/PGW-C, thus making effective use of resources.

With the introduction of CUPS, SGW-C/PGW-C have an architecture that does not connect to eNB/gNB. This is advantageous because it eliminates the need to consider the transmission distance or the number of eNBs/gNBs that can be connected. Also, since geographical considerations are no longer necessary and centralized management can be realized, maintenance efficiency can be expected to increase with a reduction in the number of maintenance sites.

SGW-U/PGW-U can be deployed according to the requirements of the services to be accommodated and their locations. For example, this will realize the concepts of deploying a large number

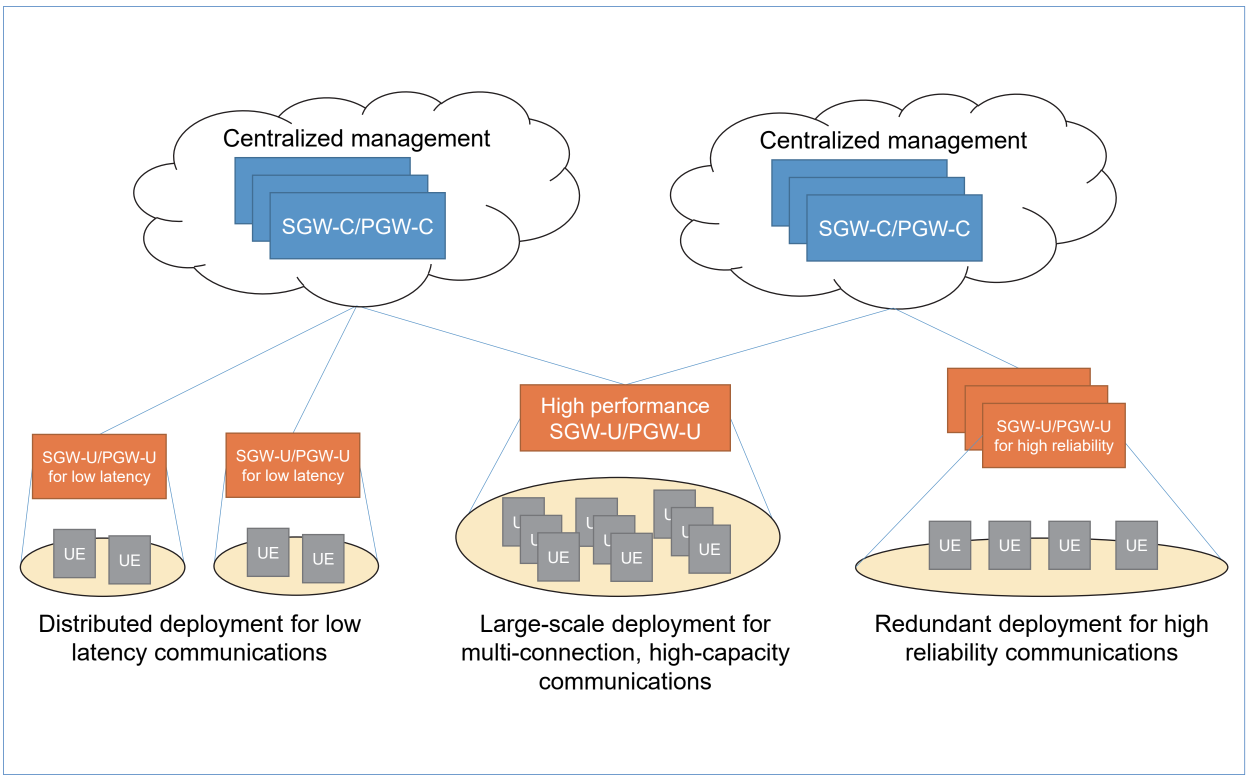


Figure 3 Example of device deployment after CUPS installation

of devices in detail nationwide for low latency data communication services, deploying high performance SGW-U/PGW-U in urban areas where there is a large population and high speed and large capacity communications are required based on regional characteristics, and deploying devices with high redundancy and reliability for disaster resilient services based on service requirements (Fig. 3).

3) Improving the Efficiency of Traffic Routes

In this section, we explain the concept and issues of traffic route efficiency before the introduction of CUPS, and how these issues can be solved with CUPS.

- Issues before the introduction of CUPS

An example of a traffic route before the introduction of CUPS is shown in **Figure 4**. Generally, in Evolved Packet System (EPS)*35, multiple sessions are set up using a multiple Packet Data Network (PDN)*36 per User Equipment (UE). An example is the use of IP Multimedia Subsystem (IMS)*37 sessions for voice services and sessions for data communication services. The two axes of traffic route efficiency are the accommodation of facilities according to service requirements and the reduction of transmission paths.

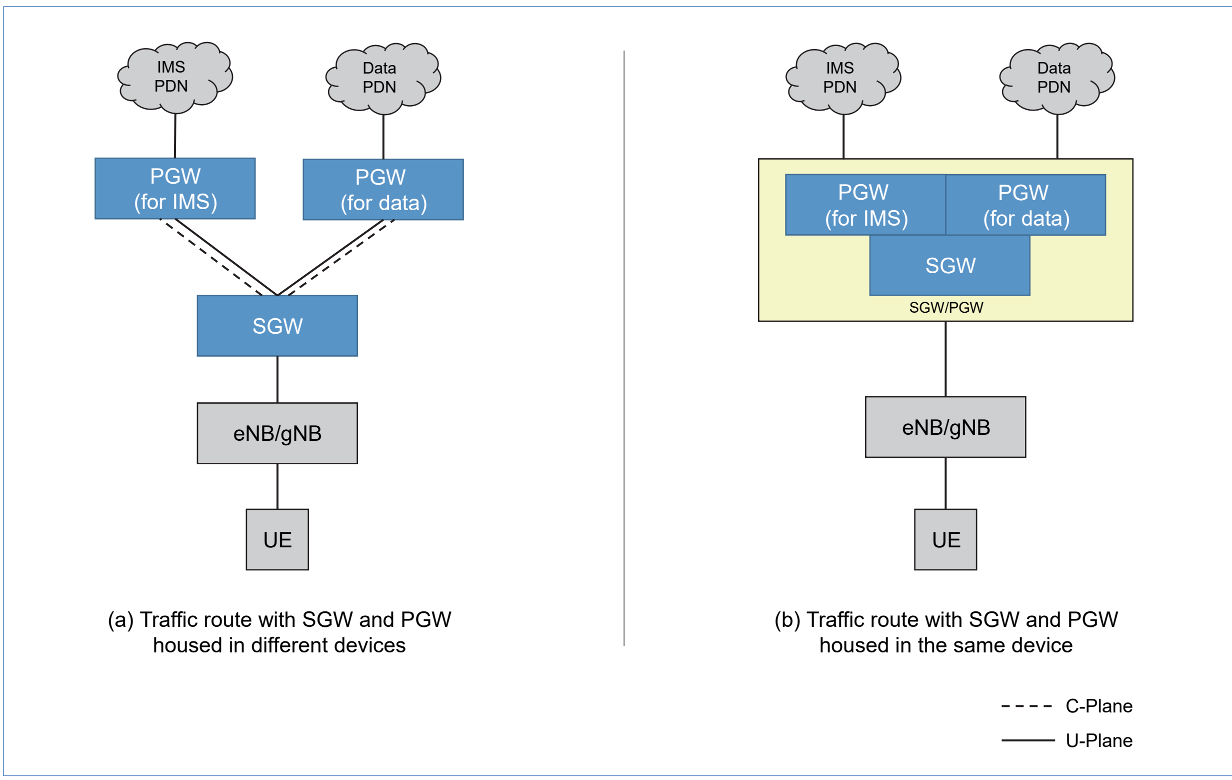


Figure 4 Example of a traffic route before the introduction of CUPS

*35 EPS: The generic name for IP-based packet networks specified by 3GPP for LTE and other access technologies.

*36 PDN: The external packet network to which the mobile core network connects.

*37 IMS: A standardized system for providing multimedia services, including voice communications, over packet communications networks.

When selecting a device according to service requirements, only one SGW per UE can be selected (Fig. 4 (a)), although it was possible to select several different PGWs according to the session with APN as the key. This led to the issue of the same SGW device requirements being applied even though service requirements differ from session to session.

Also, when considering transmission path reduction, it is possible to reduce the path of the S5 reference point by selecting SGW and PGW to be the same device by the MME. However, since only one SGW can be selected, these must be housed in the same SGW/PGW for S5 route reduction for voice and data communication services (Fig. 4 (b)). Therefore, even though PGWs can be housed in different devices depending on service requirements, the need to house them in the same device was an issue.

Thus, before the introduction of CUPS, it was difficult to both accommodate equipment to meet service requirements and reduce the transmission path.

- Improving traffic route efficiency after introducing CUPS

An example of a traffic route after the introduction of CUPS is shown in **Figure 5**. After the introduction of CUPS, SGW-U/PGW-U can be selected for each session, enabling the construction of facilities to meet service requirements.

In cases where SGW-C and PGW-C are selected from different components, each function selects U components based on independent selection rules (Fig. 5 (a)). However, since the current standard specification stipulates that APNs cannot be used when selecting SGW-U, there is a concern that traffic routes cannot be constructed using appropriate devices. For example, to simultaneously provide a voice service with high reliability requirements and a data communications service with low latency requirements to a single UE, it would be ideal to construct a route by accommodating the voice service in a device with high reliability and the low latency data communication service in a device with a short transmission distance. However, if the SGW-U cannot be selected according to the APN, there is a concern that the SGW-U for the voice service will be used to build the traffic route for the low latency data communication service.

This issue can be solved by selecting SGW-C and PGW-C as Combined SGW-C/PGW-C. SGW-C/PGW-C are deployed centrally because geographical considerations are not required, and selection is made by the MME so that SGW-C and PGW-C become Combined SGW-C/PGW-C (Fig. 5 (b)). When selected, SGW-U/PGW-U become controllable as a Combined SGW-U/PGW-U, and PGW-U selection rules can thus be applied to SGW-U/PGW-U selection parameters. As

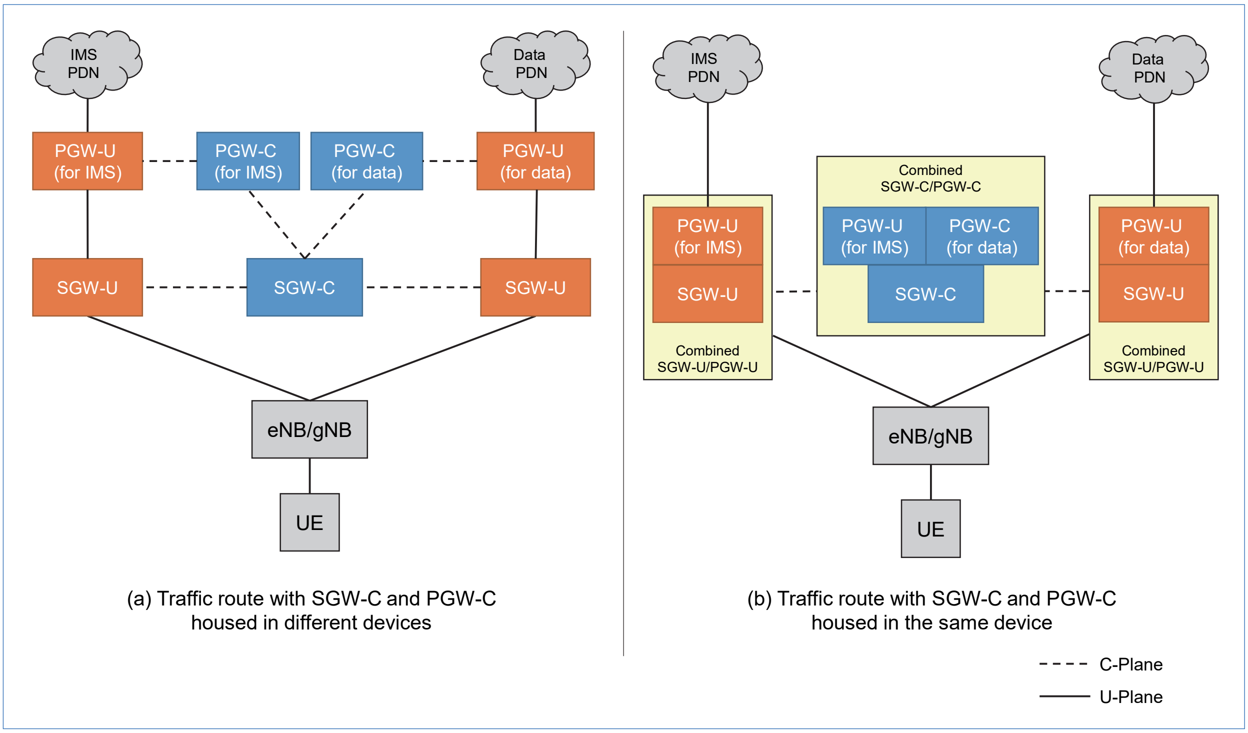


Figure 5　Example of a traffic route after the introduction of CUPS

a result, the SGW-U is selected according to the APN. From the transmission path perspective, the path of the S5-U reference point can be reduced in Combined SGW-U/PGW-U in the same way as before the introduction of CUPS. Furthermore, placing SGW-U/PGW-U closer to the eNB/gNB makes it possible to reduce the transmission distance of the S1-U reference point.

3.2 Expansion of SGW-U/PGW-U

1) Core Network Development with CUPS

SGW/PGW faces various network nodes such as eNB/gNB, MME and PCRF. Increasing the number of SGW/PGW variations was difficult because of the wide range of technical considerations to oppose existing devices. After the introduction of CUPS, SGW-U/PGW-U counterparts will be more limited than before the introduction of CUPS because the effects related to C-Plane can be closed by SGW-C/PGW-C. This will make it easier to increase the number of variations of SGW-U/PGW-U. If the new SGW-U/PGW-U to be introduced can be controlled from existing SGW-C/PGW-C, the only other consideration is the U-Plane, which has the advantage of a lowered threshold for introduction. In addition, development of SGW-U/PGW-U itself can be expected. For example, it will be possible to provide SGW-U/PGW-U to meet required service levels and equipment requirements, such as products suitable for simultaneous connection of many terminals for event venues and low-cost products for distributed deployment for low latency services. Operators will be able to develop flexible core networks by deploying SGW-U/PGW-U to suit their use cases.

2) Precautions for Device Expansion

To expand SGW-U/PGW-U, it is necessary to consider interconnectivity with SGW-C/PGW-C. PFCP clearly specifies division of functions between SGW-C/PGW-C and SGW-U/PGW-U. However, some functions may be implemented by both SGW-C/PGW-C and SGW-U/PGW-U. Therefore, it is necessary for operators to evaluate the method to be adopted.

We explain end markers as an example that needs to be evaluated in interconnectivity. The end marker is a function that notifies the end of forwarded packets to the old route when the route is switched due to handover. In the standard specification, the end marker is defined to be generated by SGW-C/PGW-C. Generation by SGW-U/PGW-U is specified as an option. However, comparing the two, the latter generation by SGW-U/PGW-U is considered to be more suitable. In terms of the amount of signal, the method generated by SGW-C/PGW-C requires an end marker transmission procedure separate from the handover control, while the method generated by SGW-U/PGW-U does not increase the amount of signal because the end marker transmission direction is performed in the handover control. In addition, in the SGW-C/PGW-C scheme, an independent session for end marker transfer between SGW-C/PGW-C and SGW-U/PGW-U needs to be established, and technical studies on the interconnection of independent sessions are required. For these reasons, we believe that the specifications generated by SGW-U/PGW-U are superior for the end marker. As such, there are cases where optional methods are used, hence, care is required when expanding the variations of SGW-U/PGW-U.

In Release 17, the functions that can be performed by both SGW-C/PGW-C and SGW-U/PGW-U, including the aforementioned end marker, were again discussed and recommended methods were clarified. NTT DOCOMO plans to follow the recommended regulations and expand the variations of SGW-U/PGW-U.

4. Development towards 5G Interworking

The 5th Generation Core network (5GC)*38 that is currently being introduced for the commercialization of the 5G SA system uses the CUPS architecture as same as in EPC. In interconnection with EPC, the SMF+PGW-C, a single device combining the 5GC Session Management Function (SMF)*39 and PGW-C has the PGW-C function of EPC. In the same way, the UPF+PGW-U, a single device combining the 5GC User Plane Function (UPF)*40 and PGW-U has the PGW-U function of EPC (**Figure 6**) [7].

SMF+PGW-C has a high degree of similarity to

*38 5GC: The core network specified by 3GPP for fifth-generation mobile telecommunications systems.

*39 SMF: The functional section that manages sessions in 5GC. Equivalent to SGW-C/PGW-C in EPC.

*40 UPF: The functional section that relays and terminates the U-Plane in the 5GC. Equivalent to SGW-U/PGW-U in EPC.

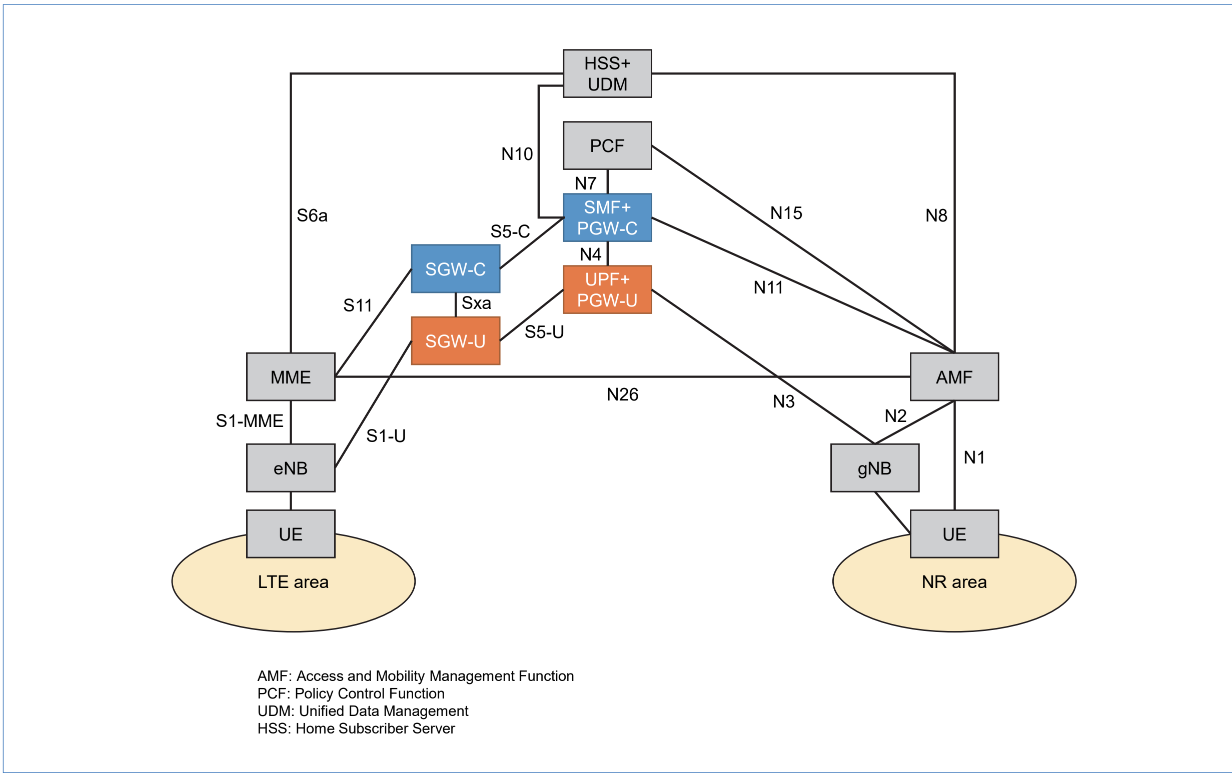


Figure 6 Introduction of SMF+PGW-C and UPF+PGW-U in EPC-5GC interwork

EPC CUPS architecture since it uses the aforementioned PFCP to control UPF+PGW-U. If CUPS architecture can be applied to EPC prior to the introduction of 5GC, a seamless transition from EPC to 5GC can be expected. Specifically, if SGW-C and SGW-U functions can be equipped in SMF+PGW-C and UPF+PGW-U, respectively, a transition plan can be considered in which either or both SGW-C and SGW-U are rolled into 5GC in stages.

5. Conclusion

This article described CUPS architecture in mobile core networks, the control scheme in CUPS architecture, the control protocol PFCP, and the GW selection scheme for flexibly selecting various SGW-U/PGW-U according to use case, and introduced the advantages of CUPS for both mobile operators and users.

NTT DOCOMO has been applying CUPS architecture to EPC to flexibly and optimally accommodate the increased traffic resulting from the introduction of 5G NSA and achieve smooth interconnection with the soon-to-be-introduced 5GC. In the future, we plan to expand the variation of SGW-U/PGW-U devices and introduce and deploy 5GC networks.

REFERENCES

[1] Y. Sagae et al.: "5G Network," NTT DOCOMO Technical Journal, Vol.22, No.2, pp.23–39, Oct. 2020.

[2] E. Endo et al.: "Overview of 5G Pre-commercial Service," NTT DOCOMO Technical Journal, Vol.21, No.3, pp.4–8, Jan. 2020.

[3] Y. Kojo et al.: "Overview of 5G Commercial Service," NTT DOCOMO Technical Journal, Vol.22, No.1, pp.4–9, Jul. 2020.

[4] 3GPP TS23.401 V15.12.0: "General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access," Sep. 2020.

[5] 3GPP TS23.214 V15.5.0: "Architecture enhancements for control and user plane separation of EPC nodes; Stage 2," Dec. 2018.

[6] 3GPP TS29.244 V15.10.0: "Interface between the Control Plane and the User Plane nodes," Sep. 2020.

[7] 3GPP TS23.501 V15.12.0: "System architecture for the 5G System (5GS)," Dec. 2020.

2021 IEEE Communications Society Best Tutorial Paper Award

On May 17, 2021, the Institute of Electrical and Electronics Engineers (IEEE) Communications Society announced the 2021 IEEE Communications Society Paper Award. Anass Benjebbour of the 5G & IoT Business Department received the IEEE Communications Society Best Tutorial Paper Award for “5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice.” This is the first of these awards for NTT DOCOMO.

The IEEE Communications Society Best Tutorial Paper Award is given to the best tutorial paper published in an IEEE journal in the past five years.

This award-winning paper was co-authored* in 2017, when the standardization of the 5th Generation mobile communication system (5G) was nearing completion, and is a comprehensive and easy-to-understand summary of the efforts made to date by industry experts from around the world. Anass Benjebbour from NTT DOCOMO has led standardization organizations such as 3rd Generation Partnership Project (3GPP)/International Telecommunication Union-Radio communication sector (ITU-R), and has summarized descriptions of the requirements for 5G, the elemental technologies to achieve those requirements, and the content and results of the demonstration experiments that NTT DOCOMO has conducted with various vendors around the world.

The paper has also been cited more than 1,000 times in just four years since it was published in the IEEE Journal on Selected Areas in Communications, Vol. 35, No. 6, in June 2017. The paper has been praised for its high citation rate and its contribution to the development of the mobile communication field and the understanding of the global technical community about 5G, which led to this award.

We will continue to contribute to mobile technology and the development of industry as a global standardization leader that promotes collaboration with various industry players in various fields.

* Co-authors: Mansoor Shafi, Andreas F. Molisch, Peter J. Smith, Thomas Haustein, Peiying Zhu, Prasan De Silva, Fredrik Tufvesson, Anass Benjebbour, Gerhard Wunder

NTT DOCOMO
Technical Journal Vol.23 No.3

Editorship and Publication
NTT DOCOMO Technical Journal is a quarterly journal edited by NTT DOCOMO, INC. and published by The Telecommunications Association.

Editorial Correspondence
NTT DOCOMO Technical Journal Editorial Office
R&D Strategy Department
NTT DOCOMO, INC.
Sanno Park Tower
2-11-1, Nagata-cho, Chiyoda-ku, Tokyo 100-6150, Japan
e-mail: dtj@nttdocomo.com